

MONTHLY NOTICES
OF THE
ROYAL ASTRONOMICAL SOCIETY
CONTAINING
PAPERS, ABSTRACTS OF PAPERS, AND REPORTS
OF THE PROCEEDINGS OF THE SOCIETY
1944

VOL. 104

Published and Sold by the
ROYAL ASTRONOMICAL SOCIETY
BURLINGTON HOUSE
LONDON, W. 1
1944

Price £1, 16s.

Printed in Great Britain by
NEILL & Co., LTD., EDINBURGH

MONTHLY NOTICES

OF THE

ROYAL ASTRONOMICAL SOCIETY

Vol. 104

No. 1

The Council elected the following as Fellows of the Society at the Council meeting on 1944 February 11:—

Frederick Walter Bowley, 50 Clifden Road, Twickenham, Middlesex;
Gordon Ellis, Blodfryn, Brynpydew, Llandudno Junction, N. Wales;
Wilfred Hutchings, The Royal School, Armagh, N. Ireland;
John Leslie Rudd, Roan, Abinger Hammer, near Dorking, Surrey; and
Arthur William Wilkinson, 127 Malham Road, Forest Hill, London, S.E. 23.

The following candidates were proposed for election as Fellows of the Society, the names of the proposers from personal knowledge being appended:—

Ivan Edward Allison, B.A., Dulce Domum, Peaks Lane, New Waltham, Lincolnshire
(proposed by W. T. Gayfer);
Cyril Evans, 3 Palace Road, Llandaff, Cardiff (proposed by G. H. Livens);
Clarence L. Friend, El Amigo Observatorio, Casa de los Amigos, Bear Valley Road,
Escondido, Cal., U.S.A. (proposed by A. H. Joy);
William Jackson, Little Ridge, Old Bursledon, Hampshire (proposed by F. Hoyle);
and
Vladimir Orloff, 56 Stuart Crescent, Hayes End, Middlesex (proposed by H. W. Cox).

MEETING OF 1944 MARCH 10

Professor E. A. Milne, President, in the Chair

The election by the Council of the following Fellows was duly confirmed:—

Dudley Perrin Bayley, Oban, Greenbank Road, Latchford Without, Warrington,
Lancashire;
Herbert John Fenner Burroughs, 93 Middle Crescent, Iscor Township, Transvaal,
S. Africa; and
John Gibson Kerr, 21 Castle Street, Edinburgh.

The following candidates were proposed for election as Fellows of the Society, the names of the proposers from personal knowledge being appended:—

Jack Howard Clarke, B.Sc., Assistant Physicist, Dominion Physical Laboratory,
Wellington, New Zealand (proposed by W. M. Smart);

Alexander Kirkcaldy, 45 Upton Street, Launceston, Tasmania (proposed by T. M. Cherry);
Charles Frederick Nelson Powell, B.Eng., 6 Lawn Road, Hampstead, London, N.W. 3 (proposed by F. J. Hargreaves); and
Howard F. Robertson, Princeton University, Princeton, N.J., U.S.A. (proposed by E. C. Bullard).

One hundred and two presents were announced as having been received since the last meeting, including:

Harvard College Observatory, Mimeograms, Series I, Nos. 1-3; Series II, Nos. 1-3; Series III, Nos. 1 and 2 (presented by Harvard College Observatory);
A. P. Herbert, *A Better Sky or Name this Star* (presented by the author); and
A 15-inch parabolic mirror, focal length 6 feet, formerly the property of the late Mr. C. R. d'Esterre (presented by Mr. J. H. Reynolds).

The President referred to the death of Professor H. F. Newall, and paid the following tribute (the Fellows standing):—

We have lost by death on 1944 February 22, at the age of eighty-six, one of our senior Fellows who was also an ex-President. Hugh Frank Newall was elected a Fellow in 1891, and served as President in 1907-1909. His early interests were in physics, to which he had intended to devote himself; and he was for a time a demonstrator in the Cavendish Laboratory at Cambridge under J. J. Thomson. But after his father had given him his 25-inch refractor, then considered a large telescope, to the University of Cambridge, Newall at some personal sacrifice transferred his interest to astrophysics, and began to work on the radial velocities of the stars. This resulted in his independent discovery, simultaneously with W. W. Campbell of Lick, of the binary nature of Capella, at a time when spectroscopic binaries were rarities. In 1909 he was appointed to the newly created Chair of Astrophysics at Cambridge; he filled this office until 1928. The McClean solar instruments and the Huggins telescopes came under his care. In 1911 the Solar Physics Observatory was moved from South Kensington to Cambridge, and Newall became Director.

Newall travelled widely, and led or took part in various eclipse expeditions. He was well known to astronomers of other countries, and was an Editor of the *Astrophysical Journal*. His distinguished appearance corresponded to a true distinction of mind. He was never so happy as when a new scientific topic engaged his ever boy-like enthusiasm, and, an amateur in the just sense of that word, he loved all things of good repute, not only in his own sciences, but also in geology and botany, in archæology and art, in farming and horticulture, in music and literature. He savoured the fine things of life, and in the hospitality of which his home was a centre he gave freely to others of the treasures of his experience. He was less well known to the younger generation of astronomers, but in all who served him he inspired great loyalty and affection. His knowledge of experimentation, his skill with instruments and his width of interest endeared him to the various assistant directors and members of the staff of the Solar Physics Observatory, to all of whom he was ever ready with suggestions and encouragement. Your President personally owed much to him. We all of us mourn the passing of a true natural philosopher, the representative of a type less common in this century than the last, and of a great gentleman, who had lived a full life in devotion to our subject.

The President announced that the Council had awarded the Gold Medal of the Society to Professor Otto Struve, Ph.D., Director of the Yerkes and McDonald Obser-

vatories, for his work on the observation and interpretation of the spectra of stars and nebulae.

The meeting then proceeded to a Discussion on Solar Phenomena. The speakers were Mr. H. W. Newton ("Solar Flares and Magnetic Storms", Second Paper), Dr. H. A. Brück, who gave an account of a paper by Dr. C. W. Allen ("Relation between Magnetic Storms and Solar Activity"), Mr. M. A. Ellison ("Sunspot Prominences—Some Comparisons between Disk and Limb Appearances"), and Dr. H. A. Brück ("On the Distribution of Intensity within the Solar Corona").

SOLAR FLARES AND MAGNETIC STORMS

(SECOND PAPER)

H. W. Newton

(Communicated by the Astronomer Royal)

(Received 1944 February 3)

Summary

Examination of a relationship found between intense (3+) solar flares within the central zone of the Sun's disk and subsequent geomagnetic activity has been extended to less intense flares (3 and 2). With flares of intensity 3 there is a small statistical rise of geomagnetic activity within a few days of the mean flare; in individual cases the disturbance is generally less intense and is certainly less probable than for flares 3+. There is no perceptible rise of geomagnetic activity after the mean flare of intensity 2.

Some distinctive flares 3 are tabulated for comparison with the most intense flares dealt with in the first paper.

Radio fade-out data confirm in a general way a small statistical rise of geomagnetic activity within a few days after the mean fade-out, but the results are less definitive than those from the solar flare data.

Introduction.—This paper continues the comparison between solar flares and magnetic storms given in *M.N.*, **103**, 244, 1943. The first paper was confined to the most intense flares of all (intensity 3+) which proved to be of special importance in the occurrence of great magnetic storms. The present paper continues the analysis to flares of intensity 3 and 2 in decreasing order of importance. Data of these less intense flares are almost wholly from the present 11-year solar cycle, for which collected observations are given in the *Zürich Bulletin*.

Flares of intensity 3 and 2, 1934/42, are now compared with:—

- (a) Magnetic storms recorded at Greenwich (Abinger).
- (b) Daily international magnetic character figures (De Bilt).

The superposed epoch method of analysis is used as described in the first paper. Some distinctive flares of intensity 3 are tabulated with comparative geophysical data. As independent data, radio fade-outs, which statistically are related to solar flares, will be compared with (a) and (b).

DATA

(Notes additional to the first paper, p. 245)

Solar Data.—Data of flares of intensity * 3 and 2 have been abstracted from *Bulletin for Character Figures of Solar Phenomena* available from 1934–41. Observations made in this country during 1942 have been included. There are four additional flares: 1869 April 13, 1928 January 26 and September 14 (from sources given in the first paper), and 1928 July 17. A spectroheliogram of the last flare (area about 250 millionths of the Sun's hemisphere: 31° E. of the central meridian) was kindly sent in 1932 to Greenwich by the Director of the Yerkes Observatory.

* For connotation, see *Trans. I.A.U.*, **5**, 293, 1935; **6**, 371, 1938; also remarks in first paper, *loc. cit.*, p. 249 (d).

Some discrimination was necessary in preparing the lists of flares 3 and 2. In those cases when such flares were observed at more than one observatory and different estimates of intensity given, the higher intensity estimate was adopted, if the listed times of observation indicated the better basis for the higher estimate. In the case of equal observing weights, preference was given to an estimate of 3, if a synchronous fade-out had been recorded, as an indication of a flare more probably of intensity 3 than of 2.* When the choice lay equally between adopting an estimate of 2 or 1, the flare was taken to be of the lower intensity. In the analysis which follows, no flare of intensity 3 was included if it occurred within 2.0 days of a flare of intensity 3+, and no flare 2 was included if likewise it fell within 2.0 days of one of higher intensity. It may be noted that the yearly frequency of flares appears to follow approximately that of the sunspots in their 11-year cycle. At minimum, flares 3+, if not flares 3, appear to be non-existent, and flares even of the lowest intensity (=1) are very infrequent, as compared with their occurrence in years near solar maximum. The total number of flares 3 thus included for discussion was 118, of which 76 were in the central zone of the disk, longitudes 0° – 45° from the central meridian. The number of flares 2 within the central zone was 332. Because of the absence of any appreciable relationship herein found between flares 2 in the central zone and subsequent magnetic activity, the analysis was not continued to flares 2 in longitudes 45° – 90° from the central meridian nor to flares of intensity 1.

Included in Table III are the areas of flares when obtainable from spectroheliograms or from drawings made with the spectrohelioscope. Mr. Evershed kindly lent an *Ha* spectroheliogram of flare No. 2 taken at Ewhurst, and Mr. Newbegin supplied drawings from his spectrohelioscope observations made at Worthing of Nos. 6 and 8.

Magnetic Data.—The magnetic storm data from Abinger are published annually in *The Observatory*. An omission from these lists is a small storm beginning 1937 June 27 having a range of 205 γ in H. The total number of storms 1934–42 inclusive is 148, of which 22 were "great" storms.

Daily international magnetic character figures are available from De Bilt from 1906 to 1939 September. For the years 1940 to 1942 the comparable magnetic character figure *K* for Abinger alone has been used in this discussion. For the derivation of *K* see *Terrestrial Magnetism*, 44, 411, 1939.

Radio Data.—Since 1936 Cable and Wireless Ltd. have reported quarterly to the Royal Observatory, Greenwich, the times of fade-outs on short-wave transmission. Although these data cover approximately only the Greenwich daylight hours (fade-outs not being recorded on channels wholly within the Earth's night hemisphere) they form a useful homogeneous body of material for a preliminary comparison with the above magnetic data (a) and (b). (Data of radio fade-outs from collective international observations, published under the auspices of the Union Radio Scientifique Internationale, are not available after August 1939.)

RESULTS AND DISCUSSION

Solar Flares 3+, 3 and 2 and Geomagnetic Activity: a Comparison

A comparison given in the first paper between flares of intensity 3+ and magnetic storms showed that in a high percentage of cases a great magnetic storm occurred within 2.0 days after an intense flare, generally when not farther than 45° from the centre of the disk. Table I shows the percentage of central zone flares of the three intensity classes associated with a magnetic storm, great or small, on one of the days 0^d , $+1^d$, $+2^d$ where 0^d =day of flare. In those cases when there was more than one storm within this time limit, only one was counted, a great storm taking precedence over a small storm. The data for flares 3+ cover the period 1859–1942 as in the first paper. The

* *M.N.*, 97, 605, 1937.

fifth column of Table I gives the approximate number of associations between flares and magnetic storms to be expected from pure chance, computed as on p. 252 of the first paper.

TABLE I
Solar Flares (Central Zone) and Associated Magnetic Storms

Solar Flare		Magnetic Storm on 0^d , $+1^d$ or $+2^d$ (0^d = day of flare)			
Intensity	Number	Observed Number	Percentage Association	Computed Number	Ratio Great/All Storms
3 +	28	23	82	4	0.70
3	77	24	31	11	0.46
2	332	62	19	48	0.13

The sudden drop in the percentage association as from flares 3 + to flares 3 is marked, as well as the drop in the relative number of great storms. For flares of intensity 2,

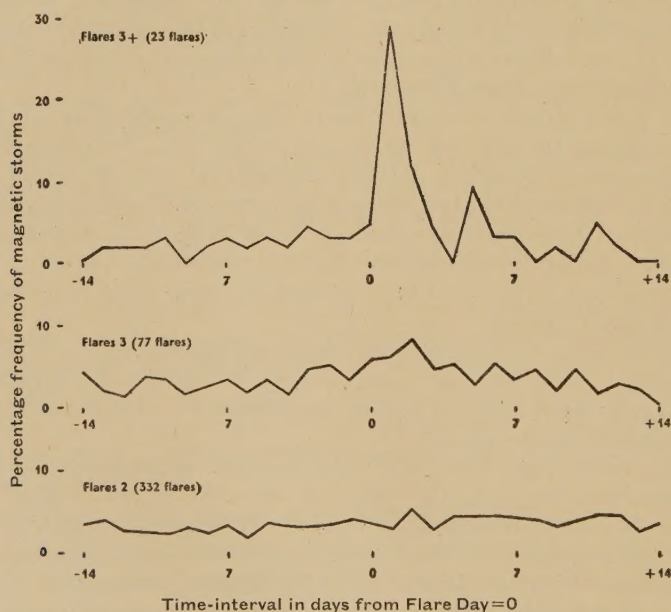


FIG. 1A.—*Solar flares (central zone) and magnetic storm frequency.*

Data: flares 3 +, 1908-42; flares 3, 1869, 1928, 1934-42;
flares 2, 1934-42.

less than two out of ten flares are associated with a magnetic storm on one of the days 0, +1, +2, a result not significantly in excess of that expected from pure chance. This excess of 14 over the computed number of chance associations is reduced to 8, if each magnetic storm is allowed to count once in only one solar flare association.

Figs. 1A and 1B show the results from the superposed-epoch method of analysis using (a) magnetic storm data, (b) magnetic character figures. Separate graphs are given for the three intensity classes of flares 3 +, 3 and 2 for the central zone (0° - 45°) of the disk.

The two sets of graphs show clearly how the marked peak at $+1^d$ or $+2^d$ given by flares of intensity 3 + collapses as we proceed to flares of decreasing intensity. A hump centred at about $+2^d$ is recognisable for flares of intensity 3, but this has disappeared in the case of flares 2.

In comparing Figs. 1A and 1B, it may be remarked that the broader peak given by the magnetic character figures may be largely ascribed to a duration effect of magnetic storms. The *commencement* day of the storm is the basis of the data in Fig. 1A. It is instructive to express the mean magnetic character figures of Fig. 1B in terms of symbols used in a Bartels' time-pattern diagram.*

If the day of normal disturbance be taken as .65 (the corresponding symbol in the diagram being a blank), then the peak of geomagnetic activity following the mean flare of intensity 3+ within the central zone will be represented by the fourth of seven black symbols used for equal steps of increasing magnetic activity. In the case of flares 3 the hump centred at about +2^d represents a mean change of little more than one symbol unit, namely, from a blank to the smallest black dot. For flares 2 there is no change of symbol throughout the whole interval from -14^d to +14^d.

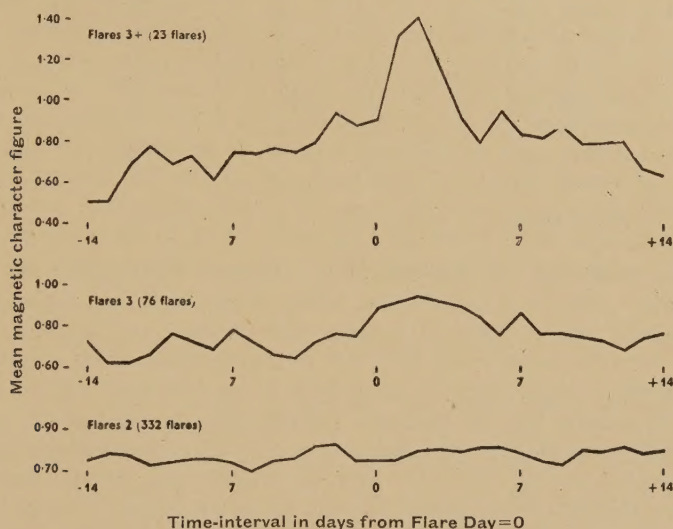


FIG. 1B.—Solar flares (central zone) and magnetic character figures.

Data : flares 3+, 1908-42; flares 3, 1928, 1934-42;
flares 2, 1934-42.

In interpreting these statistical results it will be appreciated that the smaller peak given by flares 3 as compared with that from flares 3+ may be due to (1) a smaller individual magnetic disturbance associated with the less intense flares, and/or (2) a smaller probability that a large magnetic disturbance will occur.† As seen, however, from Table I, and Fig. 1A (where great storms are given double weight), not only is there a lesser probability of a magnetic storm occurring after a flare 3 than after a flare 3+, but such storms as do occur are more generally small storms.

Solar Flares 3

(1) *Subdivision of Data.*—(a) As already shown, flares 3 in the central zone of the Sun's disk are associated statistically with a small but definite rise of geomagnetic activity centred about +2^d. In Table II, col. 2, the daily mean values of the magnetic character figures given by the superposed epoch method and graphed in Fig. 1B are grouped

* *Terrestrial Magnetism*, 37, 42, 1932.

† See Chree and Stagg, *Phil. Trans.*, 227, 33, 1928, where the probability factor is the main contribution to the diminishing amplitude of successive 27-day pulses from the primary pulse (expressing the 27-day recurrence tendency of magnetic activity or of magnetic "quiet").

successively in threes as listed in the first column. For comparison, mean magnetic character figures derived similarly from outer zone flares 3 are given in the fourth column of the table.

TABLE II
Solar Flares 3 and Magnetic Character Figures

Day Group	Mean Magnetic Character Figure and Residual from Adopted Mean			
	Inner Zone Flares		Outer Zone Flares	
d d				
-14 to -12	0.66	- 7	0.70	-7
-11 to - 9	.72	- 1	.74	-3
- 8 to - 6	.73	0	.81	+4
- 5 to - 3	.68	- 5	.78	+1
- 2 to 0	.81	+ 8	.83	+6
0 to + 2	.91	+18	.79	+2
+ 3 to + 5	.89	+16	.77	0
+ 6 to + 8	.80	+ 7	.79	+2
+ 9 to +11	.75	+ 2	.78	+1
+12 to +14	0.73	0	0.72	-5
No. of flares	76		42	

In so far as a flare may be the cause of geomagnetic disturbance, the latter cannot precede the flare. Also, it is unlikely that a flare would have any magnetic effect after 5 days, allowing for time of travel of a corpuscular emission and a possible continuance of disturbance for a day or two. If, then, the average be taken of the magnetic character figures respectively in cols. 2 and 4, excluding 0^d to $+2^d$ and $+3^d$ to $+5^d$ (as best representing the general level of magnetic activity near the time of the flares but not including any effect due to them), the residuals in cols. 3 and 5 should show any statistical difference between inner and outer zone data. It will be seen that the outer zone flares show no trace of the rise of geomagnetic activity centred around $+2^d$, which is shown by the inner zone flares.* The present data are thus in general accordance with the clearer differentiation of the effective central zone (45° radius) of the Sun's disk as shown by flare 3 + data in Fig. 3 of the first paper.†

(b) Restriction of the central zone flares to those within longitudes 0° - 26° east and west does not give an appreciably different graph of geomagnetic activity from that given by all the flare 3 data in zone 0° - 45° .

(c) Flares in east longitudes 0° - 45° have been tabulated separately from those in the same longitudes west of the central meridian. The west longitude flares are followed by the higher mean geomagnetic activity centred at about $+2^d$, the mean difference given by the two sets of data for days -2 to $+5$ being .15 magnetic character figure. The reality of this small statistical difference as between flares east and west of the central meridian must remain in doubt until more data become available from the next one or two 11-year cycles.

(2) *Some Special Cases.*—Table I of the first paper gave available data of flares 3 +, and certain characteristics were recognisable. The present table (Table III) gives a list of flares 3 that available observations showed had one or more distinctive features characteristic of flares 3 + as a whole.

These features include (1) long duration, (2) size, (3) maximum light intensity, (4) synchronous geomagnetic "crotchet". Additions to the list can doubtless be made

* Acknowledgment is gratefully made to the Astronomer Royal for suggesting this paragraph and the tabulation of residuals.

† See note, *Observatory*, 172.

by observatories possessing the requisite data of particular flares. In Table III, cols. 1 to 6, 8 to 10, and 14 correspond to similar columns in Table I of the first paper, where explanatory notes are given. Column 7 of the present table gives the maximum central intensity as measured at Greenwich: 1.0 = intensity of continuous spectrum 15A from the centre of *H α* . Columns 11 and 12 give the commencement time and duration of a fade-out, recorded by Cable and Wireless Ltd., synchronous with the solar flare. Column 13 gives the maximum range of a magnetic crotchet, synchronous with the fade-out, as recorded on the Abinger magnetic traces. With one exception the range is that in horizontal force.

Table III should be compared with Table I of the first paper. The measured areas of flares 3 are, on the average, much smaller than those of flares 3+. No. 14 was of considerable area, but the associated fade-out was of short duration. No pronounced crotchet appears on the Abinger traces. Flare no. 13, though not of great extent, was probably an intense one of long duration, for the associated fade-out lasted 70 minutes, and there was a pronounced synchronous crotchet.* It seems likely that this flare was really of intensity 3+. A great storm began 26 hours after the flare was first seen. The region was spectroscopically active during the disk transit, 27 flares having been observed. Flare no. 17 had the highest recorded peak intensity as yet measured at Greenwich. This flare was not, however, of long duration, of only medium extent and not far within the effective central zone of the disk. An absorption streamer of *H α* , having its point of origin at the flare, was seen to cross the disk and project beyond the limb as a prominence. A small magnetic storm began abruptly 41 hours after the beginning of the flare.

Fade-Outs and Magnetic Activity

In an earlier paper from Greenwich † a close relationship was found between intense solar flares and synchronous fade-outs of the "ionospheric irruption" type, often called "Dellinger fade-outs". The occurrence of a fade-out does not depend appreciably upon the position of a flare on the Sun's disk. In the case of geomagnetic activity, subsequent to flares 3+ and 3, the effective region of the disk, as already found, extends from the centre to about 45°. Although fade-outs may be regarded as giving a useful independent record of the occurrence of solar flares (chiefly those of intensity 3+ and 3) ‡, they do not locate the positions of the flares on the Sun's disk. This limitation must necessarily enter into the following statistical results, only about half of the fade-outs being attributable to flares in the central 45° region of the disk effective in the occurrence of geomagnetic storms. It may be objected that, owing to the necessarily broken solar records, it is not known whether fade-outs of the Dellinger type may occasionally occur without a visible solar flare. It seems unlikely, however, that a definite fade-out of the prescribed type and of medium or long duration is other than a geophysical accompaniment of a solar flare. Fade-outs, like flares, appear to follow the 11-year sunspot cycle. The number of fade-outs here included is 303, not counting multiple fade-outs which occur on 20 per cent. of the days of a fade-out.

The data for comparison were, as before: (a) magnetic storms, (b) magnetic character figures. For (b), the fade-outs were divided into three classes: long duration (> 40 minutes); medium (40 minutes to 21 minutes); and short (< 21 minutes). The results obtained from the superposed-epoch method are shown in the adjoining frequency

* *Observatory*, 63, 129, 1940.

† *M.N.*, 97, 594, 1937.

‡ Greenwich observations indicate that a fade-out (Cable and Wireless data) does not usually occur unless the maximum central intensity of the flare reaches the intensity of the continuous spectrum beyond the wings of *H α* . See Giovanelli and Higgs, *Terrestrial Magnetism*, 44, 181, 1939, where some numerical criteria are given.

TABLE
Distinctive Solar Flares

Ref. No.	Solar Flare						
	Date	Time, U.T.	Duration	Area	Longitude from C.M.	From Centre of Disk	Maximum Intensity
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
		h m	h		°	°	
1	1936 Aug. 28	09 20	2.2	~ 400	32 E.	44	1.3
2	1937 Jan. 27	10 36	> 1.1	270	22 W.	35	...
3	1937 June 7	11 37	> 2.4	650	21 E.	24	1.1
4	1937 June 12	{09 57 09 59	1.5 > 0.2		53 E. 43 E.	54 45	... 1.4
5	1937 July 26	09 45	> 0.8	...	31 E.	40	...
6	1937 Aug. 7	13 58	2.1	550	22 W.	27	...
7	1938 Apr. 4	07 56	3.3	...	30 E.	33	0.9
8	1938 May 11	15 03	1.0	575	37 E.	41	0.9
9	1939 Aug. 10	15 57	2.1	...	44 W.	44	...
10	1939 Aug. 19	13 40	> 3.6	...	4 W.	15	...
11	1939 Sept. 1	07 00	> 1.9	...	3 W.	20	...
12	1939 Sept. 8	11 20	> 1.2	630	26 E.	32	1.4
13	1940 Mar. 23	11 30	> 1.0	500	37 E.	40	1.1
14	1941 June 6	16 41	> 0.7	1500	34 E.	35	...
15	1941 June 8	08 57	1.1	...	12 E.	18	...
16	1941 July 3	15 09	5.5	...	4 E.	10	...
17	1942 Mar. 3	11 19	0.7	350	42 W.	44	2.6
18	1942 July 9	10 38	0.8	300	19 W.	21	1.2

Notes : Col. 4, no. 9. Exceptionally extensive.

polygon diagrams. In constructing these, days were grouped in threes, as in column 1 of Table II. The graphs show the following features:—

(1) *Fade-Outs and Magnetic Storm Frequency* (Fig. 2A).—During the interval -14^d to $+14^d$, where 0^d = day of fade-out, there is a broad maximum in the frequency of magnetic storms from -2^d to $+8^d$. This maximum, centred at about $+2^d$, indicates about twice the frequency as compared with that about a fortnight before the flare.

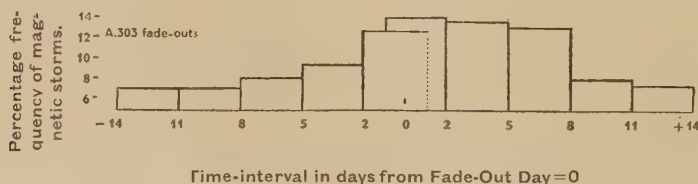


FIG. 2A.—*Fade-outs (all durations) and magnetic storm frequency (1936-42).*

(2) *Magnetic Storms and Fade-Outs* (Fig. 2B).—A maximum in the percentage frequency of fade-outs is reciprocally shown from -8^d to $+2^d$, where 0^d = day of

III

of Intensity 3 (1934-42)

Ref. No.	Associated Sunspot			Associated Geomagnetic Phenomena			
	C.M.P. U.T. (8)	Latitude (9)	Area (10)	Fade-out Begins: U.T. (11)	Duration (12)	"Crotchet" Range (13)	Magnetic Storm Within ± 2.0 Days (14)
				h m	m	'	
1	Aug. 30.0	25 S.	603	09 30	25/45	-8	None
2	Jan. 26.5	21 N.	218	10 25	85	None	None
3	June 9.1	13 N.	456	11 45	45	None	None
4	June 16.5	16 S.	897	10 00	10/50	...	None
	June 16.1	12 N.	591				
5	July 29.1	32 N.	2589	09 30	15	-9	None
6	Aug. 5.8	24 N.	634	14 05	30	-16	None
				15 40	...		
7	Apr. 6.5	23 S.	218	Partial	...	None	None
8	May 14.4	23 S.	283	15 10	...	None	May 11.7 (G)
9	Aug. 7.2	13 N.	700	15 45	45	None	Aug. 12.0 (G)
10	Aug. 19.2	8 S.	280	None	...	None	None
11	Sept. 1.2	14 S.	1800	07 30	60	None	Sept. 2.9
12	Sept 10.3	13 S.	2450	11 40	20	-28	None
				12 30	10		
13	Mar. 26.2	12 N.	1000	11 10	70	-59	Mar. 24.5 (G)]
14	June 9.5	14 N.	450	16 30	20	(-10)	None
15	June 9.5	14 N.	450	09 05	65	None	None
16	July 3.8	14 N.	500	16 20	30	-13	July 4.1 (G)
17	Feb. 28.8	...	1700	11 28	17	(+14')	Mar. 5.2
18	July 7.8	4 S.	400	10 40	50	None	None

Notes: Cols. 11 and 12, nos. 3 and 8. International Data.
Col. 13. Brackets denote that the crotchet is doubtful.

magnetic storm. This maximum, centred at about -4^d , indicates about twice the frequency of fade-outs as compared with that about a fortnight later.

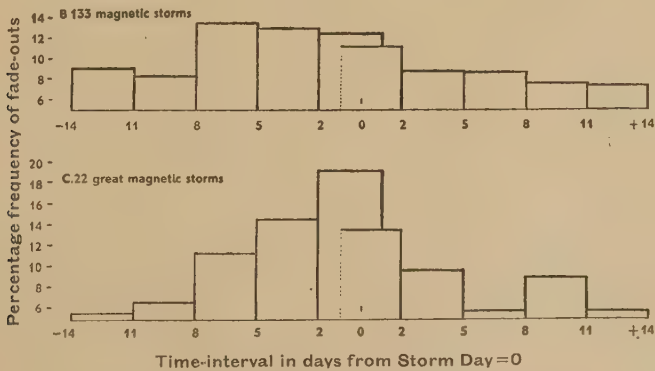


FIG. 2B.—Magnetic storms (all) and fade-outs.
FIG. 2C.—Great magnetic storms and fade-outs.

A more distinct peak at the 3-day group, -2^d to 0^d , is given by the 22 "great" storms separately (Fig. 2c). The individual daily values indicate that the peak is, in point of fact, at -1^d . Inclusion of multiple fade-outs on days adjacent to 0^d (the day of a great storm) tends to raise this peak still farther.

(3) *Fade-Outs and Magnetic Character Figures* (Fig. 3).—The mean magnetic character figure shows a small but definite rise to a broad maximum centred at

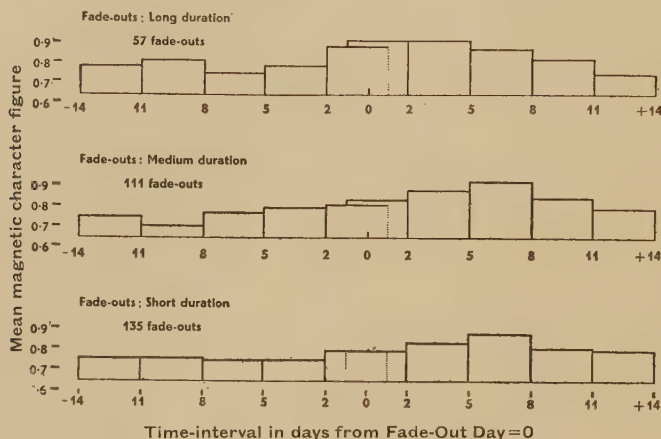


FIG. 3.—*Fade-outs and magnetic character figures.*

about $+2^d$ (long duration fade-outs), and at about $+6^d$ (medium and short duration) where 0^d = day of flare. Apart from the earlier maximum of the long duration fade-outs, there is no essential difference between the three frequency-polygon diagrams indicating a statistical rise of geomagnetic activity following fade-outs in general. It may be mentioned that the 13 fade-outs of longest duration ≥ 90 min. do not give a materially higher rise of geomagnetic activity than do fade-outs of medium or of long duration as a whole. There are, however, one or two interesting cases, *e.g.* the abnormally long duration fade-out on 1942 February 28 (lasting from $3\frac{1}{2}$ to 8 hours over some channels), synchronous with a long-continued solar flare and followed by a great magnetic storm $19\frac{1}{2}$ hours later.* But the mean maximum rise of geomagnetic activity after fade-outs is represented on a Bartels' diagram by a change of only one unit symbol.

In the preparation of this paper, material assistance was given by Miss C. Chapman of the observatory staff.

Royal Observatory,
Greenwich:
1944 January 21.

* *Observatory*, 64, 261, 1942.

RELATION BETWEEN MAGNETIC STORMS AND SOLAR ACTIVITY

C. W. Allen

(Communicated by the Director, Commonwealth Solar Observatory)

(Received 1943 December 13)

Summary

This investigation relates to the effect of individual solar phenomena on terrestrial magnetism. Magnetic disturbances have been segregated into various groups on the basis of the recurrence tendency, the storm intensity, and sudden commencements. The magnetic data have been correlated with the central meridian passage of large spot groups, with chromospheric eruptions, and ionospheric fade-outs. The method of correlation has been to plot the number of disturbances of each type occurring from five days before to fifteen days after a solar event, and to compare the counts with a random distribution. Graphs are plotted and the following rather tentative conclusions obtained:—

- (a) There is some evidence supporting the view that “great” magnetic storms are caused by chromospheric eruptions. The time interval between the eruption and the maximum of the great storm is about $1\frac{1}{2}$ days. Eruptions also cause a few smaller storms with a $2\frac{1}{2}$ -day interval.
- (b) A mean period of about 3 days is required for M-region particles to travel from Sun to Earth.
- (c) M-regions tend to avoid the area within about 40° of sunspot groups, but are more frequent than normal just outside this area.
- (d) It is *not* the eruptions within spot groups that influence M-regions.
- (e) Sunspots do not influence the commencement of an M-region, but sometimes interfere with existing ones and bring them to an end.
- (f) M-regions are identified with coronal streamers, thus explaining some characteristics of M-regions and their maximum of development two years before sunspot minimum.
- (g) Eclipse drawings show that equatorial coronal plumes are generally focussed near a sunspot, and they repel extended coronal streamers. This is thought to account for the negative sunspot–M-region correlation.
- (h) Sudden commencements are associated generally with sunspots, but are *not* due to eruptions except in the case of “great” storms.

1. *Introduction.*—It has been known for some time that magnetic storms are caused by the Sun, and that they show some statistical relation to sunspots. The most significant facts so far deduced are that great storms are associated with sunspots, but have no 27-day recurrence, while smaller storms recur but are not very definitely related to sunspots.* This seems to imply that there is some rather fundamental difference between the mechanism producing the great non-recurrent storms, and the

* Greaves and Newton, *M.N.*, **88**, 556, 1928; **89**, 641, 1929.

smaller recurrent storms. A survey of the known facts leads to a first impression that one class of storm has the following characteristics: (a) the storms are intense, (b) they are associated with sunspots, (c) they have sudden commencements, (d) they follow bright chromospheric eruptions after about a day, (e) they have no 27-day recurrence. Storms of the other class are not intense, are not related to sunspots, start gradually, do not follow chromospheric eruptions, and tend to recur after 27 days. Unfortunately it is not possible to allot individual storms with certainty to one class or the other, and there are numerous exceptions relating to all the features mentioned. It is the intention of the following work to sort out some statistical relationships connecting the above features with magnetic storms, and as far as possible to reach some conclusions concerning the mechanisms underlying the production of magnetic storms by the Sun. The investigation is related to individual storms, spots, etc., rather than the more general influence of the 11-year sunspot cycle.

2. *Magnetic Storm Data*.—The magnetic disturbances were segregated into four different groups on the basis of their recurrence tendency. Each day was regarded as a disturbed day if the international character was 1.1 or greater, or the K-sum 21 or greater. Using the 27-day charts of Bartels * the recurrent disturbed days were divided into the groups M, B and E depending on whether the disturbance was within a sequence (M for M-region), started a sequence (B for beginning), or ended a sequence (E for end). Thus the M disturbances might be taken to signify the existence of a fairly stable M-region (as defined by Bartels *) on the Sun. Storms that had relatively calm periods 27 days before and 27 days after were allotted to the non-recurrent group T (T for temporary). All disturbed days were allotted to one of the four groups M, B, E or T. A second survey of magnetic disturbances was made later, and those that fell more clearly into their group were labelled m, b, e and t, and used for a second correlation. There were 2800 disturbed days in the 37-year period 1906–1942, i.e. about 70 or 80 disturbed days per year. It will be noticed that we are dealing with much smaller disturbances than those studied by Greaves and Newton †, and also that the main subdivision of storms is related to the recurrence tendency rather than the intensity.

The period of data used is from 1906–1942. From 1906–1937 international character figures were employed, and from 1938–1942 K-index daily sums, 27-day charts were prepared for the years not included in Bartels' papers. All allotments into the groups M, B, E and T, or m, b, e and t, were based entirely on the appearance of the 27-day charts. In general, a disturbance was said to recur if a disturbance of comparable intensity occurred after an interval of 26, 27, or 28 days. In many cases it was not clear how to classify a particular day, and such doubtful cases were omitted from the m, b, e and t groups. The B, E and T types in particular were so often doubtful that the numerical strength of the b, e and t groups is only about one-third that of the B, E and T. The numbers are as follows:—

	M	m	B	b	E	e	T	t
1906–1922 . .	713	453	151	51	166	41	126	41
1923–1942 . .	977	685	240	84	273	92	154	51

Lists of sudden commencements have been compiled for the period 1923–1942 in tables similar to that published by Ogg ‡ for 1941. Only sudden commencements observed to synchronise within about 4 minutes at two or more independent observatories are regarded as genuine.

3. *Solar Data*.—For the period 1923–1942 the sunspot-number tables have been employed to compile a list of days when large spot groups passed the central meridian. Such occasions are labelled “b” in the sunspot tables. There were 267 b's in the 1923–1942 period.

* *Terr. Mag.*, 37, 1, 1932; 39, 201, 1934.

† *M.N.*, 89, 84, 1928.

‡ *Terr. Mag.*, 47, 329, 1942.

For the period 1906–1922 the Greenwich Photoheliographic Results have been used to compile a table giving the central meridian passage of spot groups whose maximum area was 400 or more millionths of the Sun's visible hemisphere. There were 374 such groups during 1906–1922. 133 larger and more complex groups (maximum area about 800 millionths) were labelled for a second correlation; of these 45 developed away from the central meridian. The "b's" in the sunspot-number tables refer to the meridian passage of spots with an area of about 600 millionths.

Bright chromospheric eruption data are obtained from the International Bulletin of Character Figures. Eruptions of importance 1 have no apparent significance, while those of importance 2 show a small correlation with terrestrial factors. Only the eruptions of importance 3 present an appreciable correlation with magnetic activity. A list has been compiled of days during 1936–1940 on which a bright chromospheric eruption (or eruptions) of importance 3 occurred.

In order to supplement the data for bright chromospheric eruptions a list of sudden radio fade-outs was prepared for the period 1936–1940. Data published by Dellinger *, Newton and Barton †, and Seaton ‡, and unpublished Mt. Stromlo data have been used to compile a list of 138 days from June 1936 to March 1940 when significant fade-outs occurred. It is well known * that all true fade-outs are caused by and are synchronous with chromospheric eruptions.

4. *Magnetic Disturbances and Meridian Passage of Large Spot Groups.*—Fig. 1 shows the number of magnetically disturbed days of the four different classes that occurred from 5 days before to 15 days after the meridian passage of a large sunspot group. The diagram can also be interpreted as the number of large spot groups passing the meridian from 15 days before to 5 days after a disturbed day.

The two periods 1906–1922 and 1923–1942 were treated independently, and some estimation of the reality of the variation can be judged by the agreement between the two curves. Thus in the case of M disturbances there appears with reasonable certainty a maximum about the day of spot passage, a minimum 3 days later, and a maximum about 8 days later. For the B disturbances there is an uncertain maximum 2 days after spot passage; the minimum at +11 days is probably fortuitous. For E disturbances there is a fairly definite maximum from 1 to 6 days after spot passage. For T disturbances there is a fairly definite maximum 4 days after spot passage.

In Fig. 2 the same data are averaged for the whole period, and expressed as a percentage of the random number of disturbances for the same period. There has been a certain amount of weighting to obtain the curves of Fig. 2. Double weight is given to m, b, e, t disturbances, and in the period 1906–1922 double weight is given to the very large spot groups. The curves do not necessarily approach 100 per cent. at each end because the relation between the number of storms and the phase of the sunspot cycle would affect the general level of the curves. For studying the effects of individual sunspots it is only the variations of the curves that are of importance.

In order to check the reality of the variations the limits of the standard deviation (*i.e.* $\pm (\text{mean per cent.})/\sqrt{n}$ where n is the mean number of storms) are shown as broken lines. By random probability laws about 1/6 of the points should lie above and 1/6 below these limits. In the B curve, for example, 4 of the 21 points lie above the limit and 3 below. This agrees well with the random law and makes it quite uncertain whether the B variation is real. In the case of the M disturbances, however, there are 10 points above the limit and 5 points below. Furthermore, the magnitude of the variation in comparison with the standard variation and the smooth nature of the curve makes its reality exceedingly probable. It will be noticed that there were about twice as many M disturbances on a day of meridian passage as 3 days after it. It seems reasonably safe to accept the reality of the main features in the E and T curves, namely the maxima near +3 or +4 days.

* *Nat. Bur. St. J. Res.*, 19, 111, 1937. † *M.N.*, 97, 594, 1937. ‡ *Terr. Mag.*, 43, 463, 1938.

When all disturbances are grouped together a very featureless curve is obtained as is seen in Fig. 2. From a statistical point of view the vague maximum from -1 to $+5$ days must be regarded as real, but the nature of the curve makes it easy to understand why no definite relation between sunspots and minor storms has been deduced up to the present.

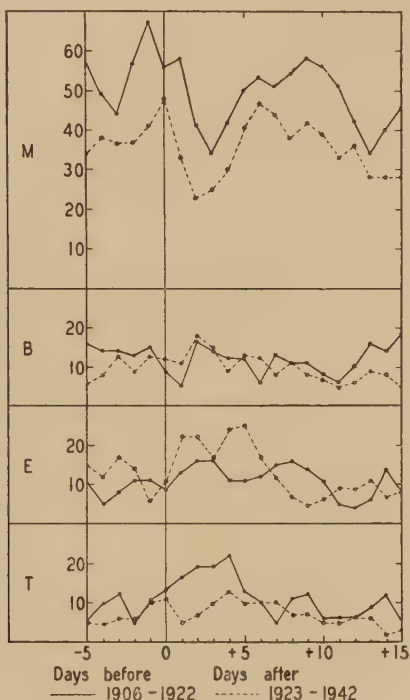


FIG. 1.

FIG. 1.—Number of disturbances of types M, B, E and T before and after meridian passage of large spot groups.

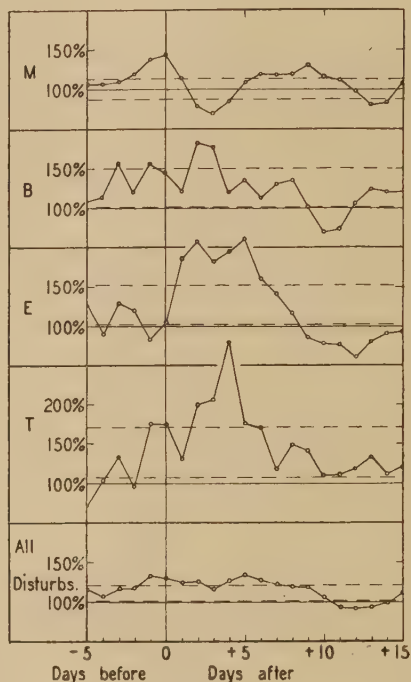


FIG. 2.

FIG. 2.—Percentage number of disturbances of various types in comparison with a random distribution before and after meridian passage of large spot groups in the period 1906-1942.

5. *The Period January 1936 to March 1940.*—A special study of this period is possible since eruption and radio data are available for comparison with the magnetic disturbances. For this period the following lists were prepared:—

- (a) List of 14 days with great magnetic storms (K index up to $7\frac{1}{2}$ or more).
- (b) 46 " with magnetic storms (K index up to 6 or more).
- (c) 32 " with magnetic storms but not great storms.
- (d) 98 " with sudden commencements.
- (e) 55 " with sudden commencements but not followed by magnetic storms.
- (f) 154 " with m magnetic disturbances.
- (g) 71 " " B " "
- (h) 81 " " E " "
- (i) 53 " " T " "
- (j) 95 " when large spot groups passed the central meridian (marked b in the sunspot tables).
- (k) 101 " on which an eruption of importance 3 was observed.
- (l) 137 " when significant fade-outs were observed.

The numbers of each type of magnetic disturbance occurring from 5 days before to 15 days after the meridian passage of large spot groups are shown in Fig. 3. In Fig. 4 the same magnetic data are plotted against the occurrence of eruptions of importance 3, and in Fig. 5 against the occurrence of fade-outs. In each case numbers above the random value are filled in to demonstrate the existence of a positive correlation. From the diagrams we see at once that sunspots, bright eruptions, and radio fade-outs all correlate positively with magnetic storms, whether great or not, and with sudden commencements. The previously determined relations between spot passage and the M, B, E and T types of disturbance are confirmed to some extent, although for the T type the data are too

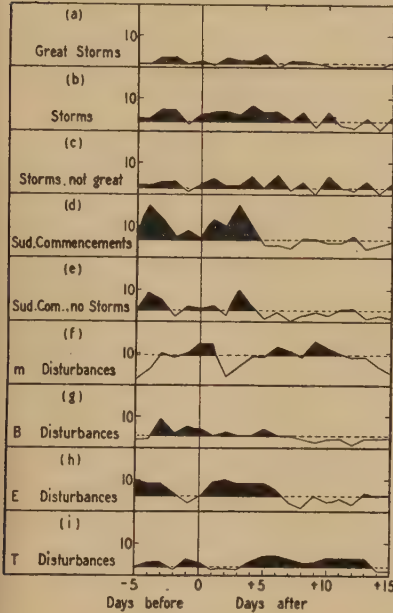


FIG. 3.

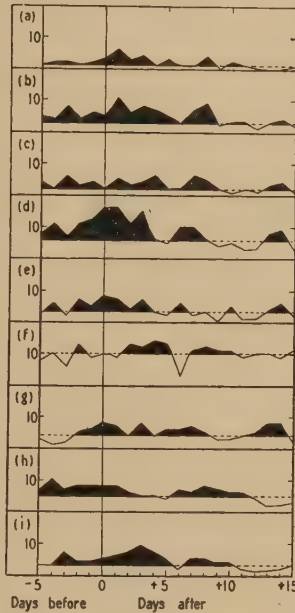


FIG. 4.

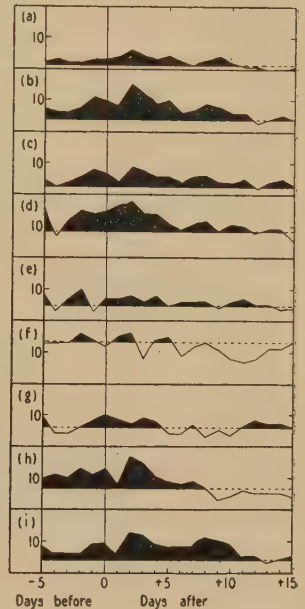


FIG. 5.

FIG. 3.—Number of magnetic disturbances of various types before and after meridian passage of large spots in the period 1936–1940.

FIG. 4.—Number of magnetic disturbances of various types before and after the occurrence of eruptions of importance during 1936–1940.

FIG. 5.—Number of magnetic disturbances of various types before and after the occurrence of radio fade-outs during 1936–1940.

meagre to give a reliable curve. The similarity between Figs. 4 and 5 is to be expected since fade-outs are caused by eruptions.

6. *Bright Chromospheric Eruptions.*—It has been suspected for some time* that great magnetic storms are caused by bright chromospheric eruptions, and the data discussed by Newton† give good support to the suggestion. There is not enough eruption data yet to prove or disprove that every “great” magnetic storm is caused by an eruption, but Fig. 4 (a) shows an increase during the days immediately following the eruption which is quite as big as could be expected when only fourteen storms have to be explained and probably several eruptions were missed. 12 of the 14 storms were preceded within three days by eruptions of importance 3. However, as soon as smaller storms are considered the relation changes. In Fig. 4 (c) the peaks following the eruption are no greater than those before it, and it must be concluded that eruptions caused only a

* Hale, *Ap. J.*, 73, 379, 1931.

† *Observatory*, 62, 318, 1939.

small fraction of these storms. Fig. 4 (i) has a fairly definite peak at +3 days, showing that when bright eruptions cause smaller disturbances the time delay is about 3 days and the disturbances non-recurrent.

If we can conclude that most of the great magnetic storms are caused by eruptions it would simplify our problems if we could show that all magnetic disturbance-sunspot correlations are due to the effect of eruptions occurring within the spot group. However, the evidence does not support this. Thus in Figs. 3 (d, e) and 4 (d, e) we find sudden-commencements correlate positively with both sunspots and bright eruptions. If, however, the sudden-commencements were caused by individual eruptions the maximum in Fig. 4 (d, e) should come a day or two after the eruption. Instead the maximum spreads right across the zero line. Also the curves for M-regions in Figs 3 (f) and 4 (f) should agree if the eruption is the real cause of the relation, whereas there is no similarity. It is evident that the individual sunspots have some influence on magnetic activity quite apart from the effect of the associated eruptions.

By their agreement with Fig. 4 the fade-out correlations shown in Fig. 5 give support to the conclusions just drawn. By comparing the maxima of Figs. 4 (a) and 5 (a) with Figs. 4 (i) and 5 (i) we see that the mean time delay between eruption and storm is about $1\frac{1}{2}$ days for great storms and $2\frac{1}{2}$ days for smaller non-recurrent storms. These times would reduce to about 1 day and 2 days if the commencements of the storms had been used instead of the storm days.

7. *M-regions*.—The positions of M-regions on the Sun can be located by the time of occurrence of recurrent (M-type) disturbances. The M-region will be t days past the central meridian on the day of the M disturbance if t is the number of days taken by the disturbing stream to travel from Sun to Earth. t is about 1 day for major magnetic storms, but may be several days for minor recurrent disturbances.

M disturbances are more frequent and intense during equinoctial months. Counts during 1906–1942 gave the following values for the probability of an M disturbance occurring during the different months of the year:—

Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
·10	·14	·19	·14	·12	·09	·09	·12	·14	·16	·11	·09

The variation is generally thought to show that the M-regions are located to the north and south of the Sun's equator, and thus affect the Earth most when the Earth's helio-centric latitude is greatest on March 7 and September 7. There is no similar frequency distribution for the temporary (T) disturbances.

Some properties of M-regions can be gathered by a study of the 27-day charts. Referring to the chart for 1923–1933 * we see that the M-regions sometimes persist for a year or more, and that they frequently continue through periods of complete sunspot inactivity. It is evident that spots are not necessary to their existence. Another feature of some significance is that there are nearly always two or three M-regions on the Sun at a time, and the disappearance of an M-region often synchronises with the appearance of another. This feature makes it probable that the M-region is not a small area coming into activity by some fortuitous chance like a sunspot. It is more likely to be an emission coming continuously from practically the whole of the Sun's surface and constrained to move in streams by forces in the Sun's atmosphere. It would then be the continuity of these streams that cause the persistence and changes of the recurrent magnetic storms.

That such a mechanism actually exists is supported by the eclipse photographs of the corona. Here we see streams extending far from the Sun's surface, but not apparently emanating from an isolated spot. If the M-regions are coronal streamers we would expect them to show their best development about one or two years before sunspot minimum when the corona has its maximum equatorial extension.† This is just the

* *Terr. Mag.*, 37, 1, 1932; 39, 201, 1934.

† *Handb. Ap.*, 7, 399, 1936.

phase of the sunspot cycle when M-regions are most intense and persistent. Some more features of the corona are discussed in Section 9 where reasons are given for expecting streamers to avoid areas near sunspots.

Turning now to the correlations demonstrated in Figs. 1 and 2 we find a decrease of the number of M disturbances about 3 days after the passage of spot groups, and an increase in the number of E disturbances, also 3 days from spot passage. This can be interpreted by assuming that 3 days is the time taken for M-region stream particles to pass from the Sun to Earth, and that a sunspot disturbance tends to interfere with an M-region if the M-region is within 40° (3 days) of the sunspot. This fits in well with the tentative assumption that M-regions and coronal streamers are identical. The maximum in the M curve at the 0 and 8 days (Figs. 1 and 2) must mean that the sunspot while deflecting the stream from the immediate vicinity of the spot increases its intensity in directions a little above 40° from the spot.

The indefinite nature of the B curves suggests that sunspots do not influence the primary commencement of M-regions.

8. *Sudden Commencements*.—Sudden commencements may be due to a sudden cause such as a bright eruption on the Sun, or to the Earth's sudden envelopment in a solar stream of particles with a rather hard boundary surface. We have shown in Section 6 that there is very little reason to suppose that sudden commencements are necessarily caused by eruptions, and we must conclude that they are the result of severe boundaries to the solar streams. The fact that sudden commencements are associated statistically with sunspots implies that the sunspot disturbance in its action of deflecting the coronal stream leaves it with a hard edge which is not natural to a stream that has formed freely. Thus sudden commencements would be more frequent among the more changing M-regions. During the period 1923–1942 the percentage disturbances starting with sudden commencements were M 12 per cent., B 25 per cent., E 24 per cent., and T 38 per cent. The small percentage for the recurrent M-regions shows that sudden commencements are not their normal characteristic, but are usually due to the influence of a sunspot while causing a deflection or some other effect on the stream. There are, however, some cases (*e.g.* 1934 July 3 and July 30) when well-confirmed sudden commencements have occurred at times when the Sun is free of spots.

Major magnetic storms generally begin with sudden commencements. If such storms are caused by eruptions we must conclude that a cloud of particles is produced which has a hard boundary surface as it approaches the Earth.

The annual and diurnal distribution of sudden commencements* are in reasonable agreement with the above conception of the sunspot effect. The maxima in March and September are explained by the variation of the Earth's heliocentric latitude which brings the Earth more into the sunspot zones during these months. The diurnal variation would be associated in some way with the angle at which the solar particles impinge on the Earth.

9. *Coronal Features*.—The main features that can be seen in the corona during an eclipse of the Sun may be classified as "arches", "plumes" and "streamers." The arches are frequently associated with prominences and sometimes they have streamers extending from them. Plumes are normally found at the poles, but features that are rather similar to the polar plumes are found sometimes in other latitudes including the sunspot zone. These "equatorial plumes" may have some influence on the Earth's magnetism, and have been studied in relation to sunspots.

The position of sunspots and equatorial plumes for 12 eclipses are shown in Fig. 6. All spots likely to be on the Sun at the time of the eclipse and within 50° in longitude of the limb have been extracted from the Greenwich Photoheliographic Results. In Fig. 6 the spot groups have been drawn the correct size for the centre of the disk and not foreshortened. Spots on both the near and far side of the Sun are included. Any

* Rodés, *Terr. Mag.*, 37, 273, 1932.

equatorial features that look in the least like plumes on the varied coronal drawings that are available * have been drawn on the outside of the disk. The more definite plumes are drawn with thicker lines.

The twelve sketches demonstrate a fairly close relationship between sunspots and coronal plumes. At the same time it can be seen that the foci of the plumes do not coincide exactly with sunspots. The plumes come from areas that in general surround or are close to spot groups. The 1919 eclipse does not show this relationship, and thus proves that the large spots and plumes are not invariably connected with one another. The appearance of well-developed equatorial plumes on about half of the twelve coronæ

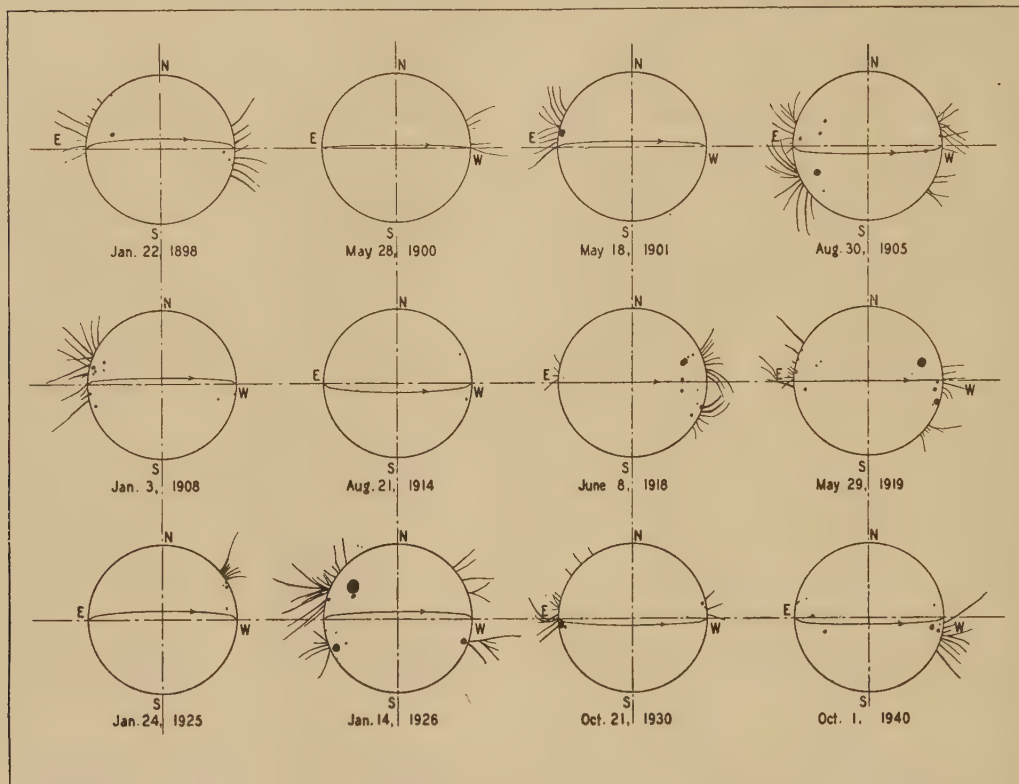


FIG. 6.

selected shows that they are much too common to be caused by eruptions in actual progress. A possibility that might be considered is that they are produced by bright hydrogen "bombs". † Such bombs are very common in the neighbourhood of active spot groups and being rather deep-seated phenomena they might produce the narrow rays seen in the plumes.

The various eclipse drawings and photographs also demonstrate that long coronal streamers are usually deflected away from such plumes, and the same is true of prominences. This would have the effect of bringing the coronal arches over the top of prominences as so often happens. The avoidance of sunspots by streamers caused by the mechanism just described should be sufficient to produce the negative correlation between sunspots and M-regions described in Section 7.

* *M.R.A.S.*, 64, App., 1927; *P.U.S.N.O.*, 4, Part 4, 1906; 10, 1926; Slocum, *Ap. J.*, 64, 145, 1926; Miller and Marriott, *P.A.S.P.*, 40, 83, 1928; Mitchell, *Ap. J.*, 71, 1, 1932.

† Ellerman, *Ap. J.*, 46, 298, 1917.

Waldmeier * has demonstrated that there is a relation between M-regions and the intensity of the coronal line near the Sun's surface. If this should prove to be a one-to-one relation some features of the mechanism here suggested will have to be modified. For example, we have visualised the base of a coronal stream as a wide area of normal solar surface, whereas Waldmeier's bright regions have a latitude extension of only 3° to 6° . The continuity of the streamers in our argument is due to regular emission from the Sun, and the main effect of solar activity is to disturb this stream. On the other hand the M-regions described by Waldmeier are obviously abnormalities. A study of the persistent M-regions associated with the present sunspot minimum should show whether Waldmeier's correlation is sufficiently close to exclude the M-region mechanism suggested in this paper.

Commonwealth Solar Observatory,

Canberra :

1943 *September 29.*

* *Z.Ap.*, 19, 21, 1939; *Observatory*, 65, 12, 1943.

SUNSPOT PROMINENCES—SOME COMPARISONS BETWEEN LIMB AND DISK APPEARANCES

M. A. Ellison

(Received 1944 January 19)

Summary

Three characteristic types of prominences associated with sunspots are considered with a view to correlating their limb and disk appearances. The discussion of the observations is carried out with special reference to: (a) type of associated spot-group, (b) form and dimensions of prominence filaments, (c) velocity and direction of motion in relation to adjacent sunspots, (d) magnetic polarity of sunspots at points of inflow.

Type IIIa.—These are inflowing filaments moving along curved trajectories (not plane figures), which originate in the surrounding chromosphere or descend from "coronal clouds" formed above the spot area. They terminate on the boundary of the penumbra with a mean velocity of inflow of 48 km./sec. Their movements seem to be independent of the magnetic polarity of the attracting sunspots. The mean projected length is 61,000 km.

Type IIIb.—Complex formations of loops and arches over spot-groups are rare. Single and double arches are frequently seen with the spectrohelioscope connecting adjacent spots of the same group. The direction of motion of the gases is upwards on one side and downwards on the other side of the arch: only one case has been observed where matter was clearly in descent on both sides of the same arch. The mean distance between the points of inflow and outflow is found to be 43,000 km., and the mean velocities at inflow and outflow +39 and -28 km./sec. respectively. The direction of motion is independent of the magnetic polarity of the spots concerned.

Type IIIc.—These are flanking prominences, often of great dimensions, the form, growth and decay of which are intimately connected with the life-history of the adjacent spot-group. They exhibit great internal activity, as well as forming external streamers which leave the tops of the prominences to flow horizontally inwards and then downwards into the spot area. After a continued existence of days, or even weeks, they may suddenly become subject to disruptive forces, rising up bodily from the chromosphere with high velocity, to descend, at least partially, in large streamers into the spot-group. An example of this phenomenon is described as it appeared on the disk.

1. *Introduction.*—With the spectrohelioscope, prominences may be observed either in *elevation* at the limb, or in *plan*, when seen as dark markings near the centre of the disk.

The disk observations possess special advantages over those made at the limb. In the first place, sunspot prominences are usually objects of low altitude, whose visibility at the limb is necessarily restricted to the short periods, usually under one day, when they are to be found in longitude 90° east or west from central meridian. They are, moreover, somewhat rare phenomena, whose incidence is confined to the more active stages in the life-history of a sunspot group. Thus the disk appearances, when these prominences are to be seen by absorption against the brighter background of the chromo-

sphere, provide us with the possibility of obtaining fuller and more continuous records, with special reference to the state of activity and stage of development of the parent spot-group. The precise location of the prominence filaments and arches in relation to the adjacent spots is also a matter of considerable importance in any discussion of the causes of their rapid movements: these measurements are more easily obtainable on the disk than at the limb.

Secondly, the Doppler shifts in the spectral lines provide us with a means of analysing the sightline * velocities and accelerations, either on the disk or at the limb, and of constructing three-dimensional pictures of the movements of each type of prominence. By this means, H. W. Newton † and the author ‡ have independently investigated prominences of *surge* type, whose movements are connected with the simultaneous appearance of a chromospheric eruption, seen by *Ha* light, upon the disk. Such records are of the highest value in supplementing the two-dimensional pictures taken with the coronagraph or spectroheliograph.

The principal difficulty met with in this work has been to correlate the limb and disk appearances in such a manner that it may be possible to assign a particular disk prominence to its corresponding limb type. Thus, for example, when a small dark flocculus makes its (often sudden) appearance near a sunspot, the observer must consider its shape, size, density and position in relation to the spot, its sightline and cross-motions, if any, and the distribution of sightline velocity along its length, before he is able to form a true picture of its character and type.

It seems desirable, therefore, to collect and classify whatever observations of this kind have become available in recent years, starting with those pioneer analyses of movements in dark flocculi, made by Hale § in 1926 when experimenting with his first spectrohelioscope. We shall follow the limb classification of sunspot prominences (Class III, *a-o*) given by E. Pettit.||

Class IIIa Prominences

2. *Historical*.—These prominences consist of brilliant jets and long fine streamers (Fig. 1, A-D), whose general appearance in elevation would suggest that they are formed of matter which is moving along lines of force in the neighbourhood of a magnetic pole.

The unique association of such prominences with sunspots, first recognised by Fényi (1891) under the term “Fleckenkronen”, has been described by him in the following terms: “Of special interest are prominences consisting of long streaks and fine threads, most of which are placed obliquely, and all converging towards a point which lies about ten thousand kilometres below the sunspot. These streaks are fugitive, often disappearing after a couple of minutes; but almost immediately others appear.” ¶ Fényi’s observations were made with an ordinary prominence spectroscope.

Similar phenomena are very fully described in the Evershed *Memoir* (1917), from records of limb prominences made with the spectroheliograph, under the heading “Jets and Rockets”.** The rockets are the long filaments, often showing great variations of brightness along their length and following trajectories which simulate the paths of rockets fired out, in all azimuths and with varying inclinations, from the spot as centre; the jets are the inner and brighter ends of these filaments, often the only parts which are visible in the spectrohelioscope. The Eversheds also describe solitary brilliant jets,

* I owe the use of this term to F. J. Sellers. When dealing with velocities, inferred from Doppler displacements observed in the chromosphere, it is less ambiguous than the term *radial-velocity* which it displaces.

† *M.N.*, 102, 2, 1942.

‡ *M.N.*, 102, 11, 1942.

§ *Ap. J.*, 71, 73, 1930.

|| *Ap. J.*, 98, 16, 1943.

¶ *P. Haynald O.*, 10, 115, 1911 (Translation by Evershed).

** *M. Kod. O.*, 1, part ii, p. 92, 1917.

"often found away from sunspots, as well as over them" *, about half of which exhibited the characteristic emission lines of "metallic" prominences. It appears very probable that many of the "solitary jets" recorded on the Kodaikanal plates should be identified with what are now termed *surges*, that is to say, thin columns of brilliant matter thrown up to heights of 30,000 km. or so before falling back along the same path to the chromosphere. Surges appear at the limb as "metallic" prominences; and, being very short-lived, they are one of the most characteristic types of chromospheric activity in the vicinity of an active spot.†



FIG. 1.—Type IIIa prominences (limb and disk).

NOTES

Limb Observations

- A. Secchi.
- B. Evershed. 1909 February 25, E. limb $+6^\circ$. Note the longest streamer apparently descending from a "coronal cloud."
- C. Sherborne. 1941 July 10, W. limb $+4^\circ$. Note the arched streamer flowing in from an adjacent spot.
- D. Sherborne. 1941 November 23, E. limb $+10^\circ$.

Disk Observations made at Sherborne

The numbers in brackets refer also to the entries in Table I.

- (37) Inflowing filament $+50$ km./sec., exhibiting "black-dot" phenomenon at point of inflow on the border of the penumbra. S, a surge of duration 23 minutes, minus followed by plus velocity.
- (38) Inflowing filaments: one shows "black-dot" phenomenon, -35 to $+47$ km./sec.
- (41) Three inflowing filaments with common velocity, 0 to $+50$ km./sec. An arch-like flocculus is descending into the following spot of the group.
- (46) A stationary flocculus feeding a downward-flowing streamer into the leading spot of the group; velocities 0 to $+35$ km./sec.

The direction of motion along the rocket-like trajectories remained in doubt until the coming of more frequent exposures by the cinematograph method. Pettit says: "In Class IIIa, in so far as appearances go, the material which forms the broken filaments seems to come from the surrounding chromosphere. Indeed, it can be assigned to no other source, since the filaments seem to be continually moving into the spot." ‡ As a result of later work, he considers that the converging streamers or broken sections of streamers "are formed in the coronal space above the spot area" §, and move down into

* *Loc. cit.*

† *Ap. J.*, 84, 333, 1936.

‡ *M.N.*, 102, 11, 1942.

§ *Ap. J.*, 98, 16, 1943.

the spot in a fan-like pattern. The spectrohelioscope records suggest that the streamers originate in both these ways, some being drawn inwards and upwards from the surrounding chromosphere and prominences before they curve over and descend steeply into the spot; others appear to descend from clouds of matter which "condense" into visibility many thousands of kilometres above the normal chromospheric level, where previously nothing luminous was seen.

3. *Spectrohelioscope Results.*—The earliest observations of such phenomena, seen in plan upon the disk, we owe to Hale. He speaks of "the rapidly changing comet-like flocculi which may form in a few seconds in the neighbourhood of active sunspots, rise with considerable velocity and pursue curved paths, at changing speeds, until their black heads impinge on the outer boundary line of the penumbra. In some cases their heads seem to reach the umbra, but the frequency with which they terminate at the apparent edge of the penumbra is remarkable. Because of their radial velocity, the line-shifter must be constantly in use to reveal them throughout their length." *

These observations are of such great interest that we have collected them into the form of a table (Table I), showing the sightline velocities of the extremities of the flocculi

TABLE I

Sightline Velocity Measures made with the Spectrohelioscope upon the Arches of Sunspot Prominences seen upon the Disk and upon inflowing Filaments

Ref. No.	Date	Time of Observation †	Sightline Velocities at Extremities of Flocculus	Probable Limb Type IIIa, b	Approximate Length of Flocculus	Type of Spot at Inflow	Lat. and Long. ‡ of Spot at Inflow	Magnetic Classification of Spot-Group §
Observer : G. E. Hale								
		P.S.T.	km./sec. → km./sec.		10 ³ km.		° °	H
1	1926 June 26	h m						
		5 21	- + 60	a	140	L	+21 139	βp, βγ 34
2	28	3 30	- + 60		67	L	+21 139	βp, βγ 34
3	Aug. 16	10 00	+ 30	a	80	L	-19 227	βp 31
4	Sept. 16	2 00	- + 30	b	31	L	+24 135	βγ, βp, βγ 38
5	17	11 16	-20 + 40	a	57	F	+23 125	" "
6	18	11 11	-40 + 26	b	100	F	+23 125	" "
7	19	10 12	+60 +110	a	61	L	+24 137	" "
8	19	10 15	+ 40	a	...	L	+24 137	" "
9	19	11 12	+14 + 54	a	72	F	+23 124	" "
10	19	11 24	+ 8 +114	a	75	F	+23 124	" "
11	19	11 55	+12 + 74	a	80	F	+23 124	" "
12	19		-32 + 34	a	56	L	+24 137	" "
13	20	9 23	- + 60	a	69	F	+23 124	" "
14	20	3 55	+ 30	a	59	L	+24 137	" "
15	21	9 00	+ 28	a	61	L	+24 137	" "
16	21	9 16	-16 + 34	a	45	F	+23 124	" "
17	21	11 23	+ 8 + 38	a	38	F	+23 124	" "
18	21	11 31	+ 8 + 18	a	47	F	+23 124	" "
19	21	4 09	0 + 27	a	52	F	+23 124	" "
20	22	9 02	- 6 + 38	a	75	F	+23 124	" "
21	22	9 48	+20 + 48	a	37	F	+23 124	" "

* M.N., 95, 468, 1935.
† The mean time for the two velocity measures.
‡ The longitudes are heliographic longitudes reckoned from Carrington's zero meridian (1854 January 1).
§ Mount Wilson classification, as published in P.A.S.P.
|| Flocculi which exhibited the "black-dot" phenomenon on the boundary of the penumbra at the point of inflow.

TABLE I—continued

Ref. No.	Date	Time of Observation *	Sightline Velocities at Extremities of Flocculus	Probable Limb Type IIIa, b	Approximate Length of Flocculus	Type of Spot at Inflow	Lat. and Long.† of Spot at Inflow	Magnetic Classification of Spot-Group ‡
Observer : M. A. Ellison								
		U.T.						H
		h m	km./sec. — km./sec.		10 ³ km.		° °	
22	1940 Aug. 2	16 40	-60 + 40		56		- 5 276	
23		19 17 20	-65		49		- 12 39	<i>lβpl</i> 32
24		30 16 25	-47 + 55	<i>b</i>	87	F	- 9 265	<i>lapd</i> 9
25	Sept. 1	11 53	-42 + 53	<i>a</i>	48	F	- 7 193	<i>lβpl</i> 31
26	Oct. 10	13 36	-43 + 37	<i>b</i>	45	L	- 15 15	<i>lβpl</i> 26
27		11 11 27	-53 + 47	<i>b</i>	28	L	- 15 15	<i>lβpl</i> 26
28	Nov. 7	14 18	-25 + 37	<i>b</i>	30	L	- 1 63	
29		14 14 49	- + 60	<i>b</i>	31	L	- 14 268	<i>dβpl</i> 20
30		27 12 00	-50 + 50	<i>b</i>	39	F	+ 10 139	<i>dβl</i> (14)
31	1941 Jan. 13	14 16	+25 + 25	<i>b</i>	50		- 3 190	<i>βd</i> 4
32	Feb. 20	13 40	-40 + 40	<i>b</i>	48	L	- 13 132	<i>lap</i> 21
33	June 17	16 06	-37 + 29	<i>b</i>	20		- 4 277	<i>dβd</i> 10
34		24 17 30	0 + 45	<i>a</i>	50	L	+ 15 217	<i>dβyl</i> 19
35	July 2	16 30	-32 + 23	<i>b</i>	81	L	+ 4 127	<i>lyl</i> 29
36	Sept. 18	11 31	-40 + 40	<i>b</i>	49	F	+ 19 208	<i>dβl</i> 23
37		30 15 41	+ 50	<i>a</i> §	31	L	+ 2 9	<i>lapl</i> 38
38	Oct. 26	14 58	-35 + 47	<i>a</i> §	77	L	+ 12 11	<i>lβpl</i> 39
39	Nov. 3	14 07	-47 + 37	<i>b</i>	81	F	+ 16 1	<i>lβpl</i> 39
40	Dec. 17	10 21	- + 46	<i>b</i>	32	F	- 5 46	<i>lβpl</i> 14
41	1942 Aug. 23	10 32	0 + 50	<i>a</i>	38	L	- 6 43	<i>dβpl</i> 21
42	1943 Sept. 7	14 13	-50 + 50	<i>b</i>	43	F	+ 14 77	<i>dβl</i> 19
43		8 11 13	- + 32	<i>b</i>	81	L	+ 15 81	" "
44		8 11 13	-30 +	<i>b</i>	59	L	+ 15 81	" "
45	Nov. 16	15 00	-32 + 41	<i>b</i>	30	F	- 8 245	
46	Dec. 20	10 30	0 + 35	<i>a</i>	50	L	- 22 138	
Observer : M. Waldmeier								
47	1936 June 12		0 + 70	<i>b</i>	26	F	+ 23 347	<i>dβfd</i> 10
48		19	0 + 30	<i>b</i>	33	F	+ 20 306	<i>dβd</i> 7
49		19	0 + 30	<i>b</i>	24	L	+ 11 308	<i>dβpl</i> 7
50	July 29		- 10 + 40	<i>b</i>	24	L	+ 26 96	<i>dβpl</i> 25
51	Aug. 17		0 + 30	<i>b</i>	12	L	+ 15 279	<i>dβpl</i> 14
52		21	-40 + 40	<i>b</i>	60	F	- 17 210	<i>dβd</i> 13
53		25	- 10 + 30	<i>b</i>	25	L	- 16 107	<i>dβpl</i> 19
54	Nov. 16		-30 + 30	<i>b</i>	13	L	- 25 164	<i>dapl</i> 14
55	Dec. 17		- 10 + 20	<i>b</i>	21	F	+ 19 72	
56		19	-25 + 25	<i>b</i>	43	L	+ 11 99	<i>dβpl</i> 17
57	1937 Jan. 26		- 10 + 30	<i>b</i>	30	F	+ 16 189	<i>dβfd</i> 22
58		27	-30 + 45	<i>b</i>	51	F	- 23 187	<i>dβpl</i> 18

NOTES

2, 22, 23, 39 were all cases of outflow, in the form of an arch, with no visible spot at the point of descent: the position given is that of the spot at outflow.

31. An arch with matter in descent on both sides, 36×10^3 km. *sp* the small group at $-3^\circ : 190^\circ$; but no spot was visible corresponding to either end of the arch. See Fig. 2.

* The mean time for the two velocity measures.

† The longitudes are heliographic longitudes reckoned from Carrington's zero meridian (1854 January 1).

‡ Mount Wilson classification, as published in *P.A.S.P.*

§ Flocculi which exhibited the "black-dot" phenomenon on the boundary of the penumbra at the point of inflow.

32, 33, 42, 43, 44. See Fig. 2. 38, 46. See Fig. 1.

37. This flocculus, 31×10^3 km. in length, moved horizontally inwards towards the spot through a distance of 86×10^3 km. in 23 minutes (average horizontal velocity of 62 km./sec.) before disappearing as a black dot at the boundary of the penumbra. See Fig. 1.

41. Three similar streamers entering a small spot from different angles with common velocity. See Fig. 1.

45. Double arch over small group of spots, both parts possessing similar velocities of ascent and descent. See Fig. 2.

1-21. Illustrations will be found in *Ap. J.*, 71, 78-95, 1930.

47-58. Illustrations will be found in *Z. Ap.*, 14, 96, 1937.

and other relevant data, together with similar observations made at Sherborne. Hale did not attempt an identification with any special type of sunspot prominence seen at the limb, but it is now reasonably clear, from his very beautiful sketches and his lucid descriptions* of the rapidly changing phenomena, that the dark flocculi to which he refers were, in fact, the inward-flowing streamers of Class IIIa prominences.

4. *Discussion of Results.*—From column 6 (Table I) we find that the average projected length of these flocculi on the disk is 61×10^3 km., and, making allowance for the curvature of the trajectory in the vertical plane, we may assume that the true length is somewhat greater than this. Twenty-two examples of similar prominences have been measured up from limb spectroheliograms, giving an average length in elevation for the curved filaments of 56×10^3 km.

The data of column 4 give us the mean sightline velocity at the point of inflow into the spot as +48 km./sec., with the greatest and least values +114 km./sec. and +18 km./sec. respectively. The outer ends of the flocculi showed values varying between -42 km./sec. (ascending) and +60 km./sec. (descending), with a mean of -1 km./sec. These figures may be compared with the velocity measures made upon 49 inward-flowing flocculi given by H. W. Newton.† From his descriptions of these flocculi it is nearly certain that they were of the same type as those we have been discussing. Newton says: "The mean inward radial velocity of these 49 flocculi at the boundary of the sunspot is 48 km./sec. The radial velocities corresponding to the far end of the flocculus show a wide range from -41 km./sec. to +46 km./sec., the mean being -3 km./sec."

These measures of velocity, considered alongside the limb pictures of IIIa filaments, combine to give us a fairly clear idea of their shape and movements. The majority are cylindrical filaments, drawn in horizontally, or nearly so, from the coronal space above the disturbed area of the spot-group, which then descend towards the penumbra with a rapidly inclining trajectory until they are moving almost vertically at the time when they impinge upon the outer boundary of the penumbra. Those which first exhibit a negative sightline velocity (motion outwards along the solar radius) are doubtless filaments drawn inwards and upwards from the surrounding chromosphere before they, like the former, curve over and descend steeply into the spot.

The "black-dot" phenomenon, which is frequently seen at the time of disappearance into the sunspot, arises from the consideration that the trajectory of the moving filament is then directed away from the observer along his line of sight, at least for those flocculi seen in the central regions of the disk. The observer is thus looking lengthwise through a long column of absorbing gas, whereas in the initial stages, when the motion is mainly crosswise to the line of sight, the depth of gas producing absorption will be no more than the normal thickness of the filament, or about 3000 km. on the average. The "contrast effect" ‡ may also operate in the same sense to increase the density of the filament at

* *Ap. J.*, 71, 73, 1930.

† *M.N.*, 94, 474, 1934. For later observations see Greenwich Photoheliographic Results (Spectroheliograph Section).

‡ *Observatory*, 64, 300, 1942.

those times when, by Doppler effect, the absorption line of the flocculus is most displaced from the centre of normal *H α* .

McMath and Pettit give height-time diagrams for the motion of IIIa knots and filaments at the limb *, from which it appears that there is an increase in velocity in most cases as the filaments descend towards the chromosphere.

It is impossible to say from the spectrohelioscope records whether there is a real increase in speed of the moving filament as the spot is approached, since the observed increase may be entirely due to the rapidly changing inclination of the trajectory during the last stages of the motion.

Another curious feature is apparent in the plan views of these flocculi. Not only are they attracted inwards and downwards towards the spot boundary, but they often partake of the general vortex motion of rotation round the spot area. Thus in a spot-group where the vortex structure is clearly developed, the inward-moving filaments seem to travel along paths which are parallel to the dark filaments outlining the vortex. The vortex filaments themselves no doubt lie at a lower chromospheric level than the former, since they stand out with greater contrast when seen in the wings of the *H α* line than at the centre of the contour.

The motion of the filaments may next be considered in relation to the magnetic polarity of the spot at the point of inflow. From the data in columns 7 and 9 we have found the magnetic polarity of the spots associated with 22 of the IIIa-type streamers; in 15 cases the streamers enter a spot of *north* polarity, in 7 cases one of *south* polarity. It appears quite certain, therefore, that the inflow is not confined to spots of one polarity.

Class IIIb Prominences

5. *General Description.*—Complex formations of loops and arches over sunspots, of the type illustrated in Fig. 2 (A–D), are of great rarity. Evershed, who gives a list of all the examples of arches known to him †, finds that they comprise only 1.6 per cent. of all sunspot prominences. They are confined to the most active stages in the life-history of very complex spot-groups. Single, or even double, arches of this type are by no means so rare, however, and many examples, seen in plan, are available for study in the spectrohelioscope records. These flocculi are composed of filaments of absorbing gas which usually rise up from the area of a newly formed spot or pore, loop over and descend into a neighbouring spot of the same group. The motion up and down the sides of these arches may be continuous for upwards of several hours.

We have here an example of the Doppler phenomenon which is of great interest and beauty. Because of the changing inclination of the trajectory to the line of sight, the arch, seen from above, is never visible in its entirety at any one setting of the spectrohelioscope line-shifter. But when the line-shifter is moved slowly back and forth about the mean position of the *H α* contour we get the characteristic “flick-over” from one side of the arch to the other, as different elements of the moving gases come progressively into view at their appropriate sightline velocities. Indeed, this remarkable effect enables us, by “hunting” with the line-shifter, instantly to distinguish one of these IIIb arches among the other types of dark flocculi surrounding a sunspot.

6. *Spectrohelioscope Observations.*—In Table I particulars relating to 31 of these arches are recorded, and 6 examples observed on the disk are illustrated in Fig. 2. The mean inward velocity has been found to be +39 km./sec. and the mean outward velocity for the other extremity of the arch –28 km./sec. The greatest and least velocities of inflow were +70 and +20 km./sec. respectively, and the corresponding values for outflow were –50 and 0 km./sec. respectively. The mean distance between the points of inflow and outflow, as given by the figures in column 6, is 43×10^3 km. There is at present no satisfactory means of estimating the true length or height of the curved

* *Ap. J.*, 88, 266, 1938.

† *Loc. cit.*, pp. 95–100.

trajectories from the disk observations. The four best-known examples of arches seen at the limb (Fig. 2, A-D) give heights varying from 40 to 100×10^3 km. The distances between the legs of the arches are evidently of the same order, though we cannot estimate these with any approach to accuracy from the limb photographs until we know the angle between the plane of the trajectory and the line of sight. The disk observations suggest that the trajectories are usually orientated approximately east and west, for example between the leading and following spots of a bipolar group.

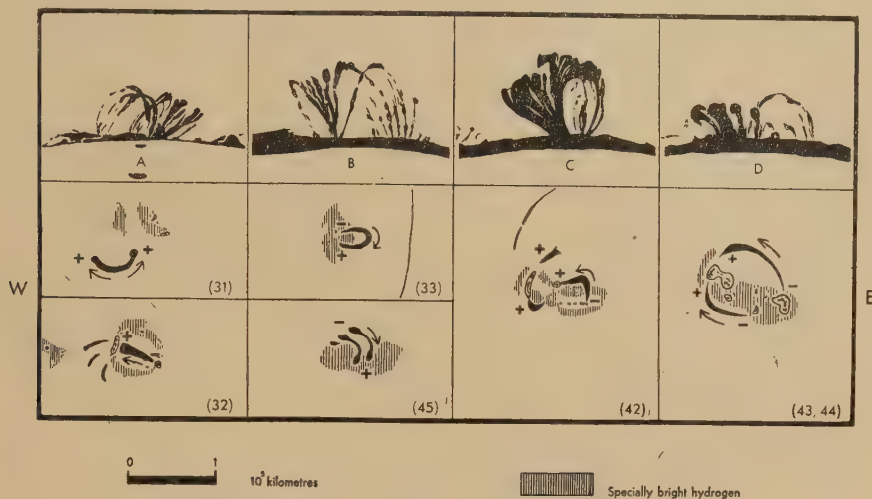


FIG. 2.—Type IIIb prominences (limb and disk).

NOTES

Limb Observations

- A. Fox, Yerkes Observatory. 1907 August 14. Standing over the spot-group, Greenwich No. 6236, E. limb.
 B. Evershed, Kodaikanal Observatory. 1910 January 7. Standing over the group, Greenwich No. 6793, W. limb, -15° . Duration > 3 hours.
 C. Mount Wilson Observatory. 1929 December 9. Standing over the group, Greenwich No. 11121, E. limb, $+6^\circ$.
 D. Mount Wilson Observatory.
 See also the arches of Fig. 3, D.

Disk Observations made at Sherborne

The numbers in brackets refer also to the entries in Table I.

- (31) An arch showing matter in descent on both sides. There were no visible spots at either extremity.
 (32) An arch with gas flowing from the following to the leading spot of the group, W. long. 47° . The other flocculi were of surge type.
 (33) Gas rising from a small spot in E. long. 60° and descending to a point south of the spot. This arch may be compared with Evershed's "Horseshoe" limb prominence, 1908 July 31.
 (45) Double arch showing identical velocities at the two ends, located over some small pores.
 (42) Gas rising from the following spot and descending into a spot in the middle of the group.
 (43, 44) Arches joining the leading and following spots of the same group one day later.

The direction of motion of the gases in 23 of these arches has been considered in relation to the magnetic polarity of the spots at the points of inflow. We find that in 10 cases the inflowing leg of the arch was located over a spot of *north* polarity; in 13 cases over a spot of *south* polarity. The direction of flow of the gases is, therefore, independent of the magnetic polarity of the spots: the same conclusion was reached by Waldmeier* for the 12 arches observed by him.

* *Z. Ap.*, 14, 91, 1937.

Only one case (No. 31) has been observed at Sherborne where matter was clearly in descent from the crown of the arch downwards along both of the legs. It thus appears that there is a considerable discrepancy between the direction of motion in the arches, as inferred from the Doppler displacements seen with the spectroheliograph, and that which has been reported by McMath from his cinematograph pictures taken at the limb. Thus McMath says: "In Class IIIb, the motion appears to be prevailing downward on both sides of the loop, there being no visible evidence of material moving up along one-half of the loop." * Pettit, on the other hand, considers that "In Class 3b the material moves along loops which rise from and return to the spot." † In this connection it may be noted that the majority of the arches seen with the spectroheliograph have been found to arise from one spot and to descend into a neighbouring spot of the same group. No case has yet been seen of a filament of gas re-entering the same spot from which it was seen previously to emerge. In a few examples only (see Notes to Table I) was there no visible spot at the point of inflow. There is evidently much scope for further investigation of these sunspot arches, with particular reference to the magnetic polarity and dipole moments of the spots concerned.

M. Waldmeier has published details with illustrations ‡ of 12 dark flocculi which were undoubtedly of the same type as those we have been discussing. We have therefore included his measures of sightline velocity in Table I. He finds that in the early stages of formation of a spot-group "the dark flocculi always flow from a younger to an older spot", a rule which has since been verified here on a number of occasions.

IIIc Prominences

8. *Characteristics of Type IIIc.*—A large bipolar spot-group seen near the limb is usually accompanied by prominences on one or both sides of it (Fig. 3, A, B). The plan views of these prominences indicate that they tend to form arcs concentric with the spot and that they often possess lateral streamers or "teeth" pointing inwards to the spot on their concave boundaries (Fig. 3, G); in other cases they may radiate from the spot area in straight lines, in a north-easterly or south-easterly direction, lengthening with the age of the group. The line-shifter of the spectroheliograph reveals that these flanking prominences are often in a state of great internal activity; but their most characteristic feature is undoubtedly the formation of streamers which leave the tops of the prominences and move horizontally inwards towards the spot area, where they curve over steeply and descend into the penumbra (Fig. 3, A-F). Bright knots and condensations follow one another along these trajectories at frequent intervals, over a period which may be many hours, and in a manner which is suggestive of the movement of conveyer boxes along a pneumatic tube. In other words, one gets the striking impression that the trajectory is a conducting path, quite distinct from the prominence material which is constrained to move along it.

The direction of motion along such trajectories was first demonstrated by Slocum §, 1910 October 8, in a series of limb photographs taken at short intervals on the IIIc prominences associated with the spot-groups, Greenwich Nos. 6893, 6894. In the longest of these trajectories (170,000 km.) condensations were found to be moving inwards towards the spot (Fig. 3, B) at speeds ranging from 16 to 60 km./sec., the speed increasing with decreasing distance from the point of inflow. The attractive force exerted by this group could be detected in prominences at a distance of 260,000 km. along the limb.

The general form of these movements has since been verified on many occasions; but it should be noted that the velocities of the condensations, as determined from distance-time plots on limb photographs, can only be approximate, since no account can be taken of the (possibly variable) sightline component. For example, a fine streamer

* *Pr. Amer. Phil. Soc.*, 79, 489, 1938.

† *Z. Ap.*, 14, 91, 1937.

‡ *Ap. J.*, 84, 333, 1936.

§ *Ap. J.*, 36, 265, 1912.

of this kind was observed with the spectrohelioscope on 1943 October 24 at the east limb (Fig. 3, F). The streamer was moving horizontally southwards from a flanking prominence and descending into a centre of attraction, located in the decaying group



FIG. 3.—Type IIIc prominences (limb and disk).

NOTES

Limb Prominences

- A. Evershed. 1912 December 24, W. limb. Sunspot at -2° indicated by arrow.
- B. Slocum. 1910 October 8, W. limb. Flanking prominences associated with groups, Greenwich Nos. 6893, 6894. Arrow indicates position of largest spot. Condensations are moving inwards along the streamer on the right.
- C. Sherborne. 1940 March 8, $11^h 20^m$, W. limb group at -5° and flanking prominence north of it.
- D. Sherborne. 1941 June 27, $17^h 30^m$. Prominences of types IIIb and IIIc associated with the spot-group at $+13^\circ$, E. limb. The IIIc prominence on the north side (left) showed great internal activity prior to the appearance of the arches.
- E. Sherborne. 1941 February 6, $13^h 30^m$ onwards. Flanking prominence feeding condensations into streamer moving southwards and descending into spot at $+11^\circ$, W. limb. Position of spot indicated by arrow. Great internal activity in prominence from $13^h 57^m$.
- F. Sherborne. 1943 October 24, $11^h 10^m$ to $15^h 00^m$, over decaying group at $+13^\circ$, E. limb. Streamer moving to the south (right) and descending into E. end of the group. Tip of streamer showed sightline velocity -45 km./sec.

Disk Observations made at Sherborne

- G. Typical IIIc flanking prominences seen upon the disk. Note the tooth-like projections facing the spot. 1940 September 27, 11^h .
- H, I, J. Stages in the disruption of an active IIIc prominence, 1941 May 4. See text. Group at $+16^\circ$, E. long. 31° .

at $+13^\circ$. Condensations could be observed moving along this trajectory southwards. The horizontal part of the trajectory showed zero sightline velocity, but the tip of the streamer where it reached the chromosphere in the vicinity of the spot had a sightline component increasing to -45 km./sec. at the point of inflow. This could only mean

that the trajectory was not confined to one vertical plane, but was being bent round horizontally as it approached the spot, possibly under the influence of a vertical magnetic field. If this is generally true of the trajectories of these inflowing streamers, it would appear to invalidate the method of deducing the velocities by measuring distances along the projected view of the trajectory: velocities so found must be subject to considerable uncertainty.

Pettit has shown * that when the streamer formation becomes well developed the flanking prominence may rise up bodily at high velocity, being torn apart into thick streamers which descend into the spot. It is not often that the destruction of such a prominence has been caught with the spectrohelioscope on the disk, though the process is evidently a frequent one, to judge by the number of occasions on which a stable dark flocculus has been seen near a spot, only to have vanished, leaving little or no trace, a few hours later.

The following example will be described for the benefit of observers who are interested in these phenomena.† On 1941 May 4, a typical IIIc prominence was seen to be in position on the disk 120,000 km. west of the group at lat. $+16^\circ$, E. long. 31° . At $9^h 18^m$ the end of the prominence nearest the spot was seen to be very dense‡, showing inward and outward movements at points A and B respectively (Fig. 3, H). The end B moved horizontally across the line of sight at an estimated speed of 80 km./sec., while also exhibiting a sightline component of approach at -70 km./sec. These velocities continued until $9^h 39^m$, by which time the foremost end of the prominence had grown out almost to the penumbra of the main spot of the group (Fig. 3, I); downward-flowing streamers then began to appear moving into the spot area at C with sightline component $+60$ km./sec. The densest part of the prominence continued to move eastward to the point D (Fig. 3, J), where at $9^h 49^m$ it was found to have a sightline component of -106 km./sec. Later this part gradually faded away while continuing its upward movement; by $10^h 16^m$ faint downward streamers could be seen returning to the chromosphere. At $10^h 43^m$ the whole area had returned to its original state, leaving only part of the original flocculus still visible west of the group.

Measures of sightline velocity (V_s) were made upon the densest, upward-moving portion of this flocculus at approximately 1-minute intervals between $9^h 26^m$ and $9^h 45^m$. By performing a graphical integration, we have deduced the value of $\int V_s \cdot dt$, thus obtaining an approximate value of 85,000 km. for the increase in height of that part of the prominence during the 19 minutes. This is quite comparable in magnitude with the upward movements recorded in similar quasi-eruptive prominences observed at the limb.

In conclusion, I wish to acknowledge my continued indebtedness to the Government Grant Committee of the Royal Society, without whose assistance these observations could not have been obtained.

* *Ap. J.*, 84, 334, 1936.

† An observation of similar type has been reported from Greenwich, *M.N.*, 91, 239, 1930.

‡ With the spectrohelioscope, a moving prominence, seen against the disk, always appears to be denser than a stationary one. This is probably due to the "contrast effect". See note, *Observatory*, 64, 300, 1942.

ON THE DISTRIBUTION OF INTENSITY WITHIN THE SOLAR CORONA

H. A. Brück

(Received 1944 February 7)

Summary

The distribution of intensity within the inner corona has been determined on a plate which had been taken during the eclipse of 1905 August 30. Values of intensity have been derived for distances between $2'.4$ and $8'.7$ from the solar limb at 72 position angles, and a series of isophotes has been drawn corresponding to a range in intensity of $2^m.5$. Intensity gradients have been determined. In the immediate vicinity of the Sun these show a definite correlation with the structure of the corona in the sense that along streamers or rays the intensity decreases less rapidly with increasing distance from the Sun than in normal regions.

1. In a paper published recently Waldmeier * has attempted a photometry of the inner parts of the solar corona which so far have been little investigated. Waldmeier's paper is chiefly concerned with the region extending from about $1'$ to $5'$ from the solar limb, but intensities are also given for slightly smaller distances and for greater distances up to about $10'$. The paper is based on a photograph taken by Wolfer during the eclipse of 1905 August 30. Among Waldmeier's results is an indication that within the inner corona intensities when measured along rays or streamers might decrease less rapidly with increasing distance from the Sun than in regions in between. The opposite effect has been observed by von Klüber † for streamers in a region which extends from about $6'$ to about $70'$ from the solar limb. Waldmeier does not, however, find any very definite correlation between intensity gradients and structure of the corona, and further observations of the intensity distribution within the inner corona appear to be desirable in view of their theoretical interest. A search at Cambridge showed that a suitable plate was available here which gives most, if not all, of the region of the inner corona studied by Waldmeier, and a photometric investigation of this plate forms the contents of the present note.

2. This photograph of the corona was taken by Newall ‡ during the same eclipse of 1905 August 30, at the same station of Guelma in Algeria, and at practically the same time, namely 16^s (time of mid-exposure) after second contact, as Wolfer's photograph. These facts make the present plate particularly interesting for a comparison with Waldmeier's measurements. The plate was obtained with a Dallmeyer telephotographic camera § with an object-glass of 4 inches aperture. The exposure time was 11^s on an Ilford Process plate which was developed in metol quinol. The scale of the present plate is somewhat larger than Wolfer's, the diameter of the Moon's image being 38.16 instead of 22.15 mm. The plate is of excellent quality and shows considerable detail within the inner corona, which can be studied up to a region close to the Moon's limb. In particular, the plate is perfectly transparent everywhere outside the corona with no detectable trace of fog. This applies also to the region inside the Moon's image except

* *Z.Ap.*, 22, 18, 1942.

† *Z.Ap.*, 2, 289, 1931. Cf. also *Z.Ap.*, 3, 142, 1931, and 4, 1, 1932.

‡ It is a pleasure to express sincerest thanks to Professor Newall for his kindness in putting this and other plates at the writer's disposal.

§ For further details cf. Professor Newall's Report, *Pr.R.S.*, A, 77, 56, 1906.

in the vicinity of the point of second contact where a faint trace of fog becomes just noticeable.

Photographic densities within the corona have been investigated with the Cambridge recording microphotometer. Records were made of the variation of density along 72 solar radii, beginning with the radius in the direction from the solar centre to the North pole of rotation and proceeding along radii with equal intervals of 5° in position angle. The tracings extended from points about 1 cm. inside the Moon's limb up to points at about the same distance outside the last traces of the corona. All records were made with a photometer slit of 0.02 mm. width and 0.5 mm. length, the slit being set at right angles to the radial direction of recording. Effects of plate grain on the tracings were negligible even for low densities. The original tracings were in excellent agreement with duplicates which, for the sake of checking the accuracy of measurement, were made for 17 different position angles.

Densities or photometer deflections were expressed in terms of the deflections produced by the region of the plate well outside the corona. These deflections agreed with those for the region inside the Moon's image, except near the point of second contact where a small correction for the presence of fog was introduced.

Each photometer record was in the first place measured at 11 points, corresponding to given distances from the solar limb. The position of the Moon's limb being well marked on the tracings, distances from the Moon's limb were measured and converted into distances from the solar limb in the usual way, using the observed values for position angles and times of second and third contact.

The points for the measurement of densities were chosen to lie between 1.1 and 1.6 solar radii from the centre of the Sun (about 1.6 to 9.5 from the limb) at equal intervals of 0.05 radii.

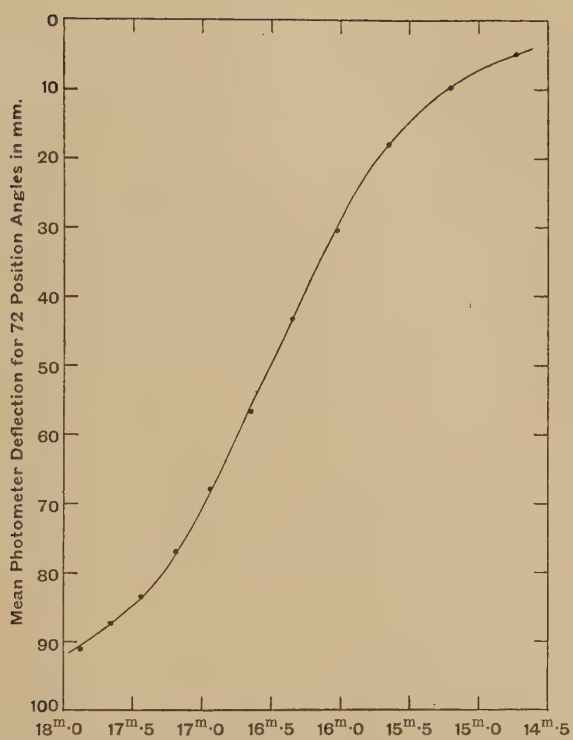
3. Like Wolfer's photograph, the present plate had not been calibrated so that photographic densities cannot directly be converted into intensities. As we are mainly concerned with a study of the deviations from the mean radial distribution of intensity, it appeared appropriate to derive a "characteristic curve" by plotting values of the mean density at given distances from the Sun and averaged over all position angles against mean values of the relative intensity derived for the same distances from other sources.* A careful study of all available observations of the mean intensity of the corona averaged over all position angles and its variation with distance from the Sun has been published by Baumbach †, whose values have been adopted in the following.

In this way the characteristic curve of Fig. 1 has been derived. Ordinates are the mean values of photometer deflection measured at the selected points (with the deflection for clear glass equalling 100 mm.), and abscissæ are the values of intensity as they follow from Baumbach's tables for the same distances from the Sun. The intensities are expressed in a scale of magnitudes with zero-point corresponding to the solar centre. The shape of the curve agrees with that which one would expect for a slow plate of high contrast. Accurate photometric measurements appear to be possible over a range of about $2^m.5$. For distances less than about 1.15 solar radii from the solar centre (2.4 from the solar limb) and greater than about 1.55 radii (8.7) photographic densities on the plate are too large or small to make accurate observations possible.

Within this range of distance (which does not extend quite so close to the Sun's limb as Waldmeier's observations) measured values of density have been converted into intensities by means of the characteristic curve of Fig. 1, and the variations of intensity with distance from the Sun have been plotted for all 72 position angles. From these

* Though both ways lead to similar results, this method has been preferred to the attempt made by Waldmeier to derive the characteristic curve from observations of the intensity distribution along the axis of rotation which Schwarzschild made during the same eclipse of 1905 (*A. Mit. Gött.*, Nr. 13, 1906).

† *A.N.*, 263, 121, 1937.



Mean Intensity of Corona at given Distances from the Sun (according to Baumbach).

FIG. 1.—Characteristic curve used for converting photographic densities into intensities.

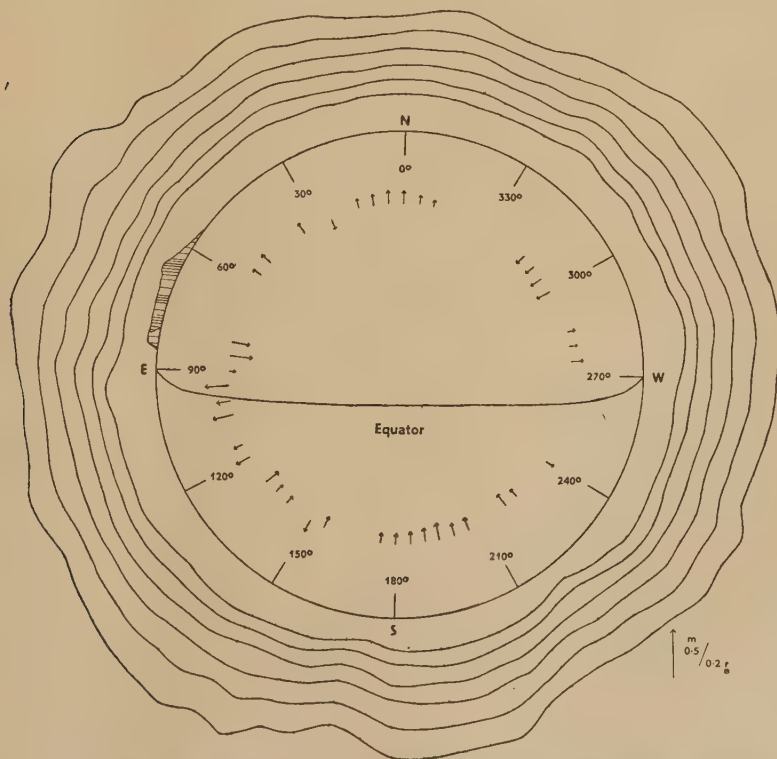


FIG. 2.—Isophotes of the corona of 1905 August 30 and residuals of intensity gradients within the inner corona.

curves distances of points showing equal intensity have been inferred, and a series of isophotes has been derived.

4. In Fig. 2 six isophotes have been drawn around the Sun, whose orientation has been indicated by marking the position of the solar equator. A conspicuous prominence at the eastern limb has also been roughly sketched in. A scale of position angles is given with its zero-point at the North pole of rotation.

Intensities differ from one isophote to the next by half a magnitude. The intensity along the innermost isophote would in Baumbach's scale be equal to $15^m.3$. The curves have been drawn through the points of observation without smoothing, since it was found that even minor local variations in the isophotes could be confirmed on the original negative. The isophotes are less smooth than those published by Waldmeier who, however, measured intensities only along half the number of radii used in the present paper.

The general shape of the corona is sensibly circular, as one expects it to be at a time close to a maximum of solar activity (sun-spot maximum at 1906.4). We can calculate the ellipticity of the corona ϵ as defined by Ludendorff*, and values of ϵ , determined for each isophote separately, are given in Table I, together with the mean diameters R of the six isophotes expressed in units of the solar diameter.

TABLE I

	<i>Ellipticities of Isophotes</i>					
Diameter R	1.16	1.22	1.29	1.37	1.45	1.56
Ellipticity ϵ	0.016	0.016	0.015	0.015	0.012	0.014

All values of ϵ are smaller than in Waldmeier's paper. When one uses Ludendorff's linear expression for the decrease of ϵ with increasing distance from the Sun—an effect which is just indicated in the figures of the Table—setting $\epsilon = a + b(R - 1)$, the following values are found for the constants a and b :

$$a = 0.017, \quad b = -0.007.$$

The value of a , which is the ellipticity of the corona in the immediate vicinity of the solar limb, is well determined. It is slightly larger than the figure derived by Ludendorff himself for the same corona, but agrees with it within the errors of Ludendorff's observations. The present value is, however, considerably smaller than Waldmeier's result $a = 0.047$, showing the effect of local differences in the shape of the isophotes on the determination of ϵ . The value of b , where $a + b$ corresponds to the ellipticity at a distance of one diameter from the Sun's centre, cannot be accurately determined from data covering only a small range of distance near the Sun.

A comparison of the isophotes in Fig. 2 with those published by Waldmeier shows general agreement on the more prominent features of the corona, but rather greater local variations in the isophotes of Fig. 2 than are shown by Waldmeier. A close inspection of the original negative and other photographs of the corona taken by Professor Newall at the same time† appears to confirm the reality of most, if not all, of these local variations in the intensity distribution of Fig. 2.

5. From the curves which give the variation of intensity with distance from the Sun for the 72 position angles used, intensity gradients were derived. When intensities were measured logarithmically those curves could be well represented by straight lines over a range of distance from the solar centre extending from 1.15 to 1.35 solar radii (2'.4 to 5'.6 from the limb). Gradients were in the first line determined for this region

* *Sber. Ak. Wiss. Berl.*, 16, 185, 1928, or cf. Dyson and Woolley, *Eclipses of the Sun and Moon*, 105, 1937.

† A fairly good picture of the corona of 1905 August 30 is given by Wesley in two drawings contained in *M.R.A.S.*, 64, App., 1927. The drawings should, however, be turned anticlockwise through about 40° if the North pole of the Sun is to be at the top of the picture.

of the inner corona, which compares with the region studied in particular by Waldmeier. Additional gradients were determined, though with rather less accuracy, for the outer region which extends from $5'.6$ to $8'.7$ from the solar limb.

Gradients have been expressed in differences of magnitude per distance of 0.2 solar radii. When gradients are averaged over all position angles their mean value for the inner corona is 1.47 ; in the same region gradients may vary with position angle by ± 0.25 . These variations are of the same order as the corresponding figures in Waldmeier's paper. For the outer region the values of the mean gradient and of the variations are 1.02 and ± 0.30 respectively.

It has been rather common to represent the variation of intensity I of the corona with distance from the Sun by an expression of the form $I = I_0/r^n$, where r is the radial distance either from the limb or the Sun's centre. The value of n in this expression may serve as a measure of the intensity gradient. Observation has shown that it is impossible to represent the intensity by a simple function of the radial distance over the whole extent of the corona, but over a limited range of distance that expression may serve as a good approximation. When applied to the present observations it leads to a mean value of $n = 8.6$ for the inner corona ($2'.4$ to $5'.6$ distance from the limb) if the distance r is reckoned from the Sun's centre; within the same region n varies with position angle between 7.2 and 9.9 .

In Fig. 2 the residuals of the gradients as previously defined and determined for the inner corona are plotted as arrows inside the solar disk. Arrows pointing outwards denote gradients larger than the average, arrows pointing inwards smaller gradients. Only residuals exceeding ± 0.07 have been plotted, the length of the arrow in the lower right-hand corner of the figure denoting a gradient of 0.5 .

A definite connection is seen between residuals in Fig. 2 and the structure of the corona. Along rays or fans and wherever the outer isophotes bulge outwards, gradients within the inner corona are below the average and are appreciably smaller than in "normal" regions in between those features. Negative residuals are evident, for example, in the broad fan near the South pole between position angles 175° and 205° , and between position angles 300° and 315° , or 130° and 140° . Gradients not necessarily negative, but small compared with neighbouring regions, are found at position angles 25° , along the streamer at 155° and at 110° .

The only definite exception from this rule is found in the region between position angles 275° and 290° , where we might expect negative instead of positive residual gradients. However, the corona has a rather more complicated structure near the solar equator than in other regions, and this may well be the cause of non-uniformity.

There appears to be sufficient evidence from these observations to support Waldmeier's original suggestion that within the inner corona the intensity falls off less rapidly with increasing distance from the Sun along streamers or rays than in normal regions. Such a result would be naturally expected if it was not for von Klüber's observations, which leave little doubt that the opposite is true for the outer corona where intensities decrease most rapidly along streamers. It would have been of interest to study up to greater distances the behaviour of those rays which are visible in the corona of 1905, but plates taken at the same time and with the same instrument, and with longer exposures which show the outer corona, are rather fogged and not suitable for photometry. Gradients determined for the outer region of our plate ($5'.6$ to $8'.7$ from the solar limb) might be expected (so far as their rather smaller accuracy goes) to show a state of transition between inner and outer corona and therefore to behave more like von Klüber's observations. No indications of such an effect is, however, indicated in our observations.

NOTE ON THE WAVE-LENGTH OF THE SODIUM LINE D_2 IN THE SUN

J. Evershed, F.R.S.

(Received 1944 February 15)

Summary

Attention is drawn to an error in the wave-length of the sodium line D_2 in the Revision of Rowland's Table, probably due to an unsuspected telluric line on the red edge of D_2 at approximate wave-length 5890.08. Accurate measures of the position of D_2 can only be made in light from the western limb of the Sun when the solar line is shifted to red and covers the telluric line.

Measures are given of the interval $D_1 - D_2$ at the western limb and in the vacuum tube, showing that in both it is 5.973 in agreement with Kiess's value in the vacuum arc. The true values of D_1 and D_2 in the Sun are deduced and compared with Kiess's values of these lines in the vacuum arc, showing that the shift towards red is the same in both lines and in close agreement with the relativity effect.

In the course of measurement of a large number of comparison spectra of the sodium lines D_1 and D_2 in the Sun and vacuum tube for estimating the value of the Einstein relativity shift, it was found that the interval $D_1 - D_2$ as given in Rowland's Table and repeated in the Revision of Rowland is incorrect and does not agree with the same interval between these lines in the vacuum tube.

In the Revision of Rowland the wave-lengths are

D_1	5895.944	Interval 5.967
D_2	5889.977	

The actual interval should be increased by 6 units in the third decimal, according to my measures. The error is in the wave-length of D_2 and is probably due to the effect on the original measures of Rowland of an unsuspected telluric line on the redward edge of D_2 , which cannot be seen in spectra of the centre of the Sun, but which broadens the line on the red side. This line, like all telluric lines, varies with the altitude of the Sun and with the humidity of the atmosphere. It can be observed separated from D_2 only when this line is shifted towards blue at the Sun's eastern equatorial limb, giving 2 km./sec. approach, and during the period including the last week of September and the first week of October, when the Earth approaches the Sun at 0.5 km./sec. Also to this may be added one or two tenths due to the rotation of the Earth when the observation is made soon after sunrise. In all, a shift of D_2 amounting to 0.05 Å. or 2.6 km./sec. is needed to separate the lines.

I have estimated the position of this telluric line to be λ 5890.08 referred to nearby water-vapour lines, as given in the Revision of Rowland, and in the few plates which show it the intensity is about 2 on Rowland's scale.

In Rowland's Table a water-vapour line of intensity 2 is given at λ 5889.888 on the blue edge of D_2 . This should appear nearly separate from D_2 at the Sun's western equatorial limb at the end of March when the Earth is receding from the Sun 0.5 km./sec., but I have failed to find evidence of it at such times. A line in the position given would widen D_2 towards the blue and make the interval $D_1 - D_2$ larger than it is in the vacuum

tube spectrum. In a later series of solar spectra taken in the spring of 1943 for finding the relativity shift of D_1 at the Sun's limb, I have in 12 plates measured the interval in west limb spectra, and find it to be 5.971 Å. or 0.002 Å. *smaller* than in the vacuum tube spectrum, doubtless due to the influence of the line on the red side of D_2 .

Assuming that the line at the blue edge of D_2 does not exist, accurate measures of the position of D_2 and of the interval $D_1 - D_2$ can only be obtained at the Sun's western limb near the solar equator, the telluric line being then completely absorbed within the solar line. Measuring by the positive on negative method on such spectra obtained with the Hargreaves liquid prisms, I have obtained very consistent results using the telluric lines as given in the table to determine the scale of the plates. These give the intervals as follows:—

	$D_1 - D_2$
7 west limb spectra . . .	$5.9733 \pm .0004$
6 vacuum tube spectra ' . . .	$5.9731 \pm .0005$

The actual distance separating the D lines in these spectra is about 9 mm. In angstroms it agrees exactly with Kiess's interferometer measures of D_1 and D_2 , which give the interval 5.973.

In order therefore to get the true value of D_2 at the centre of the Sun, I take the value of D_1 as given in the Revision of Rowland, but this should be decreased by approximately 0.004 angstrom, which is an average correction to reduce it to the I.A.U. 1928 standards. Thus, in the Revision of Rowland

	D_1 is λ 5895.944
Approximate correction	$- .004$
	5895.940
Interval as found above	5.973
	$D_2 = 5889.967$

Taking Kiess's absolute values of D_1 and D_2 we can now find the shift of the solar lines towards red at the Sun's centre.

	D_2	D_1
Centre of Sun . . .	5889.967	5895.940
Kiess vacuum arc . . .	5889.954	5895.927
Shift in Sun . . .	+ .013	+ .013
Relativity shift . . .	+ .0125	+ .0125

These values of the shift at the centre of the Sun depend to some extent on the correction applied to D_1 in Rowland's Table, but their equality proves that the interval $D_1 - D_2$ must be accurate. The shift agrees exactly with my direct measures of it in D_1 as given in *M.N.*, 98, 195 (1936).

*Highbroom,
Ewhurst, Surrey.*

THE PROBLEM OF AN ACCURATE DETERMINATION OF THE RELATIVISTIC LIGHT DEFLECTION

E. Finlay Freundlich and W. Ledermann

(Received 1943 November 23)

Summary

The following investigation discloses the characteristic features of the problem of determining the light deflection L with sufficient accuracy. The main points are:

(i) The relative weight of L is small even in the most favourable case of a symmetrical field and non-degenerate normal equations, unless an exceptionally large number of stars is observed in the immediate vicinity of the Sun. It is not sufficient to observe a few bright stars near the Sun, and the exposure times have to be chosen accordingly. For a camera of aperture 5 inches and a ratio of apertures of about $1/40$ the time of exposure should not be less than 2 minutes; the focal length itself should not be smaller than 20 to 30 feet. Cameras of such a length reduce the scattered light in the neighbourhood of the Sun sufficiently to render stars down to the 10 mag. phot. observable. The Lick star field with an exposure time of 2 minutes shows 11 stars fainter than 10 mag. situated at a distance of not more than 4.5 solar radii from the centre of the Sun. Plates obtained at the 1929 eclipse with the horizontal camera of 8 inches aperture and 27.9 feet focal length show stars 1.5 mag. fainter than do similar plates taken with the astrographic lens of equal aperture but only 11.2 feet focal length.

(ii) The inclusion of a scale correction S and of L as unknowns in the Least Square solution has to be avoided; it reduces the weight of the light deflection and brings the determinant of the normal equations critically close to its vanishing point, unless exceptionally numerous stars have been observed quite near the Sun. The scale correction has therefore to be determined separately and with such an accuracy that its standard deviation multiplied by h does not give rise to an uncertainty in L which surpasses the preassigned limit. The weight of L obtained in this case is very sensitive to any asymmetry in the star field (see (23)).

Introduction.—Since 1929 several attempts have been made to measure the light deflection in the gravitational field of the Sun and to verify the ensuing results; but all these attempts have failed. Thus the question remains unsettled whether the deflection at the Sun's limb amounts to $1''.75$ as predicted by the Theory of Relativity, or whether it is substantially larger, namely equal to $2''.2$ as indicated by the findings of the Potsdam Expedition of 1929 and confirmed by a renewed discussion of previous observational material.

Moreover, this last determination revealed certain fundamental difficulties which have made it clear that a definite decision between the values $1''.75$ and $2''.2$ is bound to constitute a formidable task. The accuracy of observation must be such that the standard deviation (mean error) of the final result does not exceed $0''.1$. This accuracy

is attainable only if the experiment is planned and carried out according to a method which has full control of the sources of errors and is able to assess in advance the expected error of the result.

The aim of the present paper is to develop the theory of observational errors of the above problem and to discuss its bearing on the design of future experiments.

1. The light deflection is determined by a differential comparison of two plates one of which, the "eclipse" plate, is exposed to the star field surrounding the Sun during a total eclipse, whilst the other, the "night" plate, is exposed to the same star field with the same camera, when the star field stands on the night sky.

Suppose the field contains n stars. The positions of the stars on the two plates are specified by their co-ordinates relative to rectangular systems which, it is assumed, may be brought to complete coincidence save for unavoidable experimental errors. Nevertheless it will be found that the position vectors of the i th star on the two plates differ by a small vector

$$(\delta x_i, \delta y_i) \quad (i = 1, 2, \dots, n), \quad (1)$$

and the problem consists in analysing this discrepancy into its significant and accidental causes. The equations used for this purpose are somewhat more general than those hitherto employed, allowance being made for an eccentric position of the Sun in the star field and on the plate, namely

$$\left. \begin{aligned} N_1 - (y_i - \beta)\Theta + (x_i - \alpha)^2 P + (x_i - \alpha)(y_i - \beta)Q + (x_i - \alpha)S + (x_i/r_i^2)L &= \delta x_i, \\ N_2 + (x_i - \alpha)\Theta + (x_i - \alpha)(y_i - \beta)P + (y_i - \beta)^2 Q + (y_i - \beta)S + (y_i/r_i^2)L &= \delta y_i \end{aligned} \right\} \quad (2)$$

($i = 1, 2, \dots, n$). The equations involve 7 unknowns, namely N_1 , N_2 , Θ , P , Q , S and L , the coefficients of which depend on the co-ordinates (x_i , y_i) of the stars on the night plate and on the co-ordinates (α , β) of the point where the optical axis intersects the plate. The origin is taken at the centre of the Sun, and

$$r_i^2 = x_i^2 + y_i^2.$$

The unknowns are of three types:

(i) L : the terms in L represent the physical effect of the light deflection for whose sake the experiment has been undertaken.

(ii) N_1 , N_2 , Θ : these quantities reflect the unavoidable small errors in matching the two plates, namely a small displacement (N_1 , N_2) of the origins and a small angle Θ between the two x -axes. These are genuine errors of observation and cannot be eliminated.

(iii) P , Q , S : the terms in P and Q , which are quadratic in x and y , enter if the plates are not strictly at right angles to the optical axis, thus causing a slight distortion of the field. The *scale factor* S takes account of a possible change in the focal length between the times of the two exposures. There is no intrinsic reason why the three unknowns of this group should not be determined separately and independently of the others, as they affect each plate individually and are not directly related to the comparison of the two plates. Indeed, our subsequent analysis will show that only by adopting such a procedure with regard to S can a satisfactory solution for L be obtained. For the sake of lucidity we shall confine ourselves to the terms in S alone and assume that P and Q have been dealt with by other methods. The essential features of the problem are not altered thereby; equations (2) now become

$$\left. \begin{aligned} N_1 - (y_i - \beta)\Theta + (x_i - \alpha)S + (x_i/r_i^2)L &= \delta x_i, \\ N_2 + (x_i - \alpha)\Theta + (y_i - \beta)S + (y_i/r_i^2)L &= \delta y_i \end{aligned} \right\} \quad (3)$$

or in matrix notation

$$\mathbf{A}\mathbf{u} = \mathbf{d}, \quad (4)$$

where

$$\mathbf{A} = \begin{bmatrix} \mathbf{I} & \mathbf{O} & -y_1 + \beta & x_1 - \alpha & x_1/r_1^2 \\ \mathbf{I} & \mathbf{O} & -y_2 + \beta & x_2 - \alpha & x_2/r_2^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{I} & \mathbf{O} & -y_n + \beta & x_n - \alpha & x_n/r_n^2 \\ \mathbf{O} & \mathbf{I} & x_1 - \alpha & y_1 - \beta & y_1/r_1^2 \\ \mathbf{O} & \mathbf{I} & x_2 - \alpha & y_2 - \beta & y_2/r_2^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{O} & \mathbf{I} & x_n - \alpha & y_n - \beta & y_n/r_n^2 \end{bmatrix} \quad (5)$$

is a matrix with $2n$ rows and 5 columns, and where

$$\mathbf{u} = \begin{bmatrix} N_1 \\ N_2 \\ \Theta \\ S \\ L \end{bmatrix} \quad \text{and} \quad \mathbf{d} = \begin{bmatrix} \delta x_1 \\ \vdots \\ \delta x_n \\ \delta y_1 \\ \vdots \\ \delta y_n \end{bmatrix} \quad (6)$$

are column vectors of 5 and $2n$ elements respectively. The "best" estimate $\hat{\mathbf{u}}$ of $\mathbf{u}$ is obtained in the usual way by passing from (4) to the *normal equations*

$$\mathbf{A}'\mathbf{A}\hat{\mathbf{u}} = \mathbf{A}'\mathbf{d}. \quad (7)$$

It is readily found that

$$\mathbf{A}'\mathbf{A} = n \begin{bmatrix} \mathbf{I} & \mathbf{O} & -\bar{y}' & \bar{x}' & \xi \\ \mathbf{O} & \mathbf{I} & \bar{x}' & \bar{y}' & \eta \\ -\bar{y}' & \bar{x}' & a' & \mathbf{O} & \beta\xi - \alpha\eta \\ \bar{x}' & \bar{y}' & \mathbf{O} & a' & \mathbf{I} - \alpha\xi - \beta\eta \\ \xi & \eta & \beta\xi - \alpha\eta & \mathbf{I} - \alpha\xi - \beta\eta & h^{-1} \end{bmatrix}, \quad (8)$$

where

$$\bar{x}' = \frac{1}{n} \sum_{i=1}^n (x_i - \alpha), \quad \bar{y}' = \frac{1}{n} \sum_{i=1}^n (y_i - \beta), \quad a' = \frac{1}{n} \sum_{i=1}^n \{(x_i - \alpha)^2 + (y_i - \beta)^2\}, \quad (9)$$

$$\xi = \frac{1}{n} \sum_{i=1}^n \frac{x_i}{r_i^2}, \quad \eta = \frac{1}{n} \sum_{i=1}^n \frac{y_i}{r_i^2}, \quad \frac{1}{h} = \frac{1}{n} \sum_{i=1}^n \frac{1}{r_i^2}. \quad (10)$$

In this paper we are not concerned with the actual values of the estimates which can be calculated from (7), but rather with the measure of precision, *i.e.* the *weights* which may be attributed to the results, especially in the case of L . The weight w is usually* defined to be inversely proportional to the square of the standard deviation σ . But as only ratios of weights will ultimately be of significance, we shall simply put

$$w = 1/\sigma^2.$$

If it is assumed that the readings

$$\delta x_1, \delta x_2, \dots, \delta x_n, \quad \delta y_1, \delta y_2, \dots, \delta y_n$$

are all of the same weight w_0 say, the weight w_L of the calculated value of L is given by†

$$w_L/w_0 = \Delta/\Delta_{LL}, \quad (11)$$

* A. C. Aitken, *Statistical Mathematics*, 108, 1942.

† Whittaker and Robinson, *The Calculus of Observations*, 241, 1924.

where Δ is the determinant of $\mathbf{A}'\mathbf{A}$ and Δ_{LL} is the minor obtained by deleting the last row and the last column of Δ . After some elementary manipulations it is found that (11) is equivalent to

$$w_L/w_0 = n \left[\left(\frac{1}{h} - \frac{1}{a} \right) - \frac{(a\xi - \bar{x})^2 + (a\eta - \bar{y})^2}{a(a - \bar{x}^2 - \bar{y}^2)} \right], \quad (12)$$

where ξ , η and h are defined in (10) and

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i, \quad \bar{y} = \frac{1}{n} \sum_{i=1}^n y_i, \quad a = \frac{1}{n} \sum_{i=1}^n r_i^2. \quad (13)$$

This result forms the basis of the subsequent discussion. We remark at once that, whether or not the field is symmetrical about the Sun, the position of the optical axis no longer enters explicitly, as the expression on the right-hand side of (12) is independent of a and β . The quantities (13) are simple functions of the co-ordinates relative to the centre of the Sun.

As was emphasized in the Introduction, the object of the experiment is to decide whether the value of L is $1''.75$ or $2''.25$. Since the difference between these values is $0''.5$, the standard deviation σ_L of L should not exceed $0''.1$, *i.e.* we ought to have

$$w_L = 1/\sigma_L^2 \geq 1/0.01 = 100. \quad (14)$$

The standard deviation of a single reading δx_i or δy_i will not be much below $0''.3$; if k plates have been obtained, the initial standard deviation is given by

$$\sigma_0 = 0.3/\sqrt{k},$$

and consequently

$$w_0 = 1/\sigma_0^2 = k/0.09.$$

Thus in order to ensure the success of any experiment it is desirable that

$$w_L/w_0 \geq 9/k.$$

It will be seen from Table I that this requirement has not yet been satisfied and would be met only by a sufficiently large value of k . This is due to the fact that the determinant Δ is almost zero, which means that the set of equations (7) is almost degenerate. The algebraical expression of this degeneracy is the smallness of the term

$$n \left(\frac{1}{h} - \frac{1}{a} \right),$$

where h^{-1} and a^{-1} are respectively the reciprocals of the *harmonic* and *arithmetic means* of the quantities $r_1^2, r_2^2, \dots, r_n^2$.

It will be seen in the next section that this critical term can be avoided by excluding S from the Least Square solution.

If the star field were *perfectly symmetrical* with respect to the Sun, we should have

$$\bar{x} = \bar{y} = \xi = \eta = 0$$

(see equations (13) and (10)), and (12) would reduce to

$$(w_L/w_0)_{\text{sym}} = n \left(\frac{1}{h} - \frac{1}{a} \right). \quad (12')$$

Since

$$a - \bar{x}^2 - \bar{y}^2 = \frac{1}{n} \sum_{i=1}^n \{(x_i - \bar{x})^2 + (y_i - \bar{y})^2\}$$

is never negative, it follows from (12) and (12') that

$$(w_L/w_0)_{\text{sym}} \geq w_L/w_0.$$

Thus the symmetrical star field represents the most favourable case. At the same time it is desirable to make the expression (12') as large as possible. Theoretically this may be accomplished by observing as many stars as possible in the immediate neighbourhood of the Sun, which will tend to increase the difference $h^{-1} - a^{-1}$. But in actual fact, the number of stars near the Sun will not suffice to make this expression large enough. This is borne out by Table I: the extremely favourable star field observed by the Lick Expedition had 43 stars within 7.5 solar radii (out of a total of 92 stars). Yet the relative weight of L is too small, and the wide extension of the field is of no avail; for a star whose distance from the Sun exceeds about 7.5 solar radii contributes practically nothing towards the solution of the problem.

2. From the preceding discussion it has become apparent that the normal equations (8) cannot yield a sufficiently accurate determination of the light deflection L , because the determinant of $\mathbf{A}'\mathbf{A}$ is too small. To obviate this difficulty the cause of the degeneracy has to be removed, and the only course open to us is to eliminate the unknown S from the Least Square solution by a separate determination. If the value of the scale factor arrived at in this way be denoted by S_0 , the equations (3) are written

$$\begin{aligned} N_1 - (y_i - \beta)\Theta + (x_i/r_i^2)L &= \delta x_i - (x_i - \alpha)S_0, \\ N_2 + (x_i - \alpha)\Theta + (y_i/r_i^2)L &= \delta y_i - (y_i - \beta)S_0 \end{aligned} \quad (15)$$

($i = 1, 2, \dots, n$), or in matrix notation

$$\mathbf{B}\mathbf{v} = \mathbf{d} - S_0\mathbf{z}, \quad (16)$$

where

$$\mathbf{B} = \begin{bmatrix} 1 & 0 & -y_1 + \beta & x_1/r_1^2 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & 0 & -y_n + \beta & x_n/r_n^2 \\ 0 & 1 & x_1 - \alpha & y_1/r_1^2 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 1 & x_n - \alpha & y_n/r_n^2 \end{bmatrix}; \quad \mathbf{v} = \begin{bmatrix} N_1 \\ N_2 \\ \Theta \\ L \end{bmatrix}; \quad \mathbf{z} = \begin{bmatrix} x_1 - \alpha \\ \vdots \\ x_n - \alpha \\ y_1 - \beta \\ \vdots \\ y_n - \beta \end{bmatrix} \quad (17)$$

The normal equations corresponding to (16) are

$$\mathbf{B}'\mathbf{B}\hat{\mathbf{v}} = \mathbf{B}'(\mathbf{d} - S_0\mathbf{z}), \quad (18)$$

where

$$\mathbf{B}'\mathbf{B} = n \begin{bmatrix} 1 & 1 & -\bar{y}' & \xi \\ 0 & 1 & \bar{x}' & \eta \\ -\bar{y}' & \bar{x}' & a' & \beta\xi - a\eta \\ \xi & \eta & \beta\xi - a\eta & h^{-1} \end{bmatrix}. \quad (19)$$

Denoting the determinant of $\mathbf{B}'\mathbf{B}$ by Δ' and that of its leading minor by Δ'_{LL} , we find that the weight of the estimate of L is now given by

$$w'_L/w_0 = \frac{\Delta'}{\Delta'_{LL}} = n \left\{ \frac{1}{h} - (\xi^2 + \eta^2) - \frac{(\bar{x}\eta - \bar{y}\xi)^2}{a - \bar{x}^2 - \bar{y}^2} \right\}, \quad (20)$$

where the quantities on the right are defined by (10) and (13). If we neglect the terms

that are due to the asymmetry of the star field, we get

$$(w'_L/w_0)_{\text{sym}} = n/h.$$

The predominant term in (20) is now nh^{-1} instead of the much smaller term $n(h^{-1} - a^{-1})$ in (12), and it is for this reason that the present method, although more elaborate, is likely to succeed where the former fails.

3. The method outlined in the preceding section presupposes an independent determination of the scale factor S . Consequently any error committed in this determination affects the value of L and that of the other unknowns. It is therefore important to know how accurately S must be determined, if the validity of the final result is not to be seriously impaired.

Consider two different estimates $S^{(1)}$ and $S^{(2)}$ of S . Each of these values gives rise to estimates of the remaining variables, which in accordance with (18) are given by

$$\begin{aligned}\mathbf{B}'\mathbf{B}\mathbf{v}^{(1)} &= \mathbf{B}'(\mathbf{d} - S^{(1)}\mathbf{z}), \\ \mathbf{B}'\mathbf{B}\mathbf{v}^{(2)} &= \mathbf{B}'(\mathbf{d} - S^{(2)}\mathbf{z}).\end{aligned}$$

Hence if we put

$$\delta\mathbf{v} = \delta\mathbf{v}^{(1)} - \delta\mathbf{v}^{(2)} = \begin{bmatrix} \delta N_1 \\ \delta N_2 \\ \delta \odot \\ \delta L \end{bmatrix} \quad (21)$$

and

$$\delta S = S^{(1)} - S^{(2)},$$

we find that

$$\mathbf{B}'\mathbf{B}\delta\mathbf{v} = -\delta S\mathbf{B}'\mathbf{z}, \quad (22)$$

where

$$\mathbf{B}'\mathbf{z} = n \begin{bmatrix} \bar{x} \\ \bar{y} \\ 0 \\ 1 - \alpha\xi - \beta\eta \end{bmatrix}.$$

Solving (22) for δL , i.e. for the last element of $\delta\mathbf{v}$, we obtain

$$\delta L = -\delta S \frac{1 - (\bar{x}\xi + \bar{y}\eta)}{\frac{1}{h} - (\xi^2 + \eta^2) - \frac{(\bar{x}\eta - \bar{y}\xi)^2}{a - \bar{x}^2 - \bar{y}^2}}. \quad (23)$$

In this case the asymmetry terms may become considerable, but if the star field be perfectly symmetrical we should have

$$\delta L = -h\delta S.$$

This shows that an initial error in S induces an error in L which is about $\bar{r}^2$ times as great. Thus the determination of S has to be carried out with the utmost accuracy.

4. The results of the foregoing theory have been applied to the observations made by the Greenwich *, Lick † and Potsdam ‡ Expeditions. The relevant data, as far as they are available to us, have been collected in the following table.

* Dyson and Eddington, *M.R.A.S.*, 62, App., 1920.

† Campbell and Trumpler, *L.O.B.*, 346, 1923.

‡ Freundlich, Klüber und v. Brunn, *Abh. Preuss. Ak. Wiss.*, 1931, 1; *Ann. Bosscha Obs. Lembang*, 5, B, 1933. Freundlich und v. Brunn, *Z. Ap.*, 6, 218, 1933. Freundlich und Gleissberg, *Publ. Istanbul Obs.*, 2, 1935.

TABLE I

Star field	Aperture	Exp. time	Limit mag.	M	k	$r_{\max}$	$\bar{r}$	$1/a$	$1/h$	w_L/w_0	w'_L/w_0	σ_L
						solar radii						
Greenwich	4", 1/57	28 ^s	6.0 phot.	7	7	5.5	3.6	1/14.5	1/7.9	0.400 *		1.6 $\times \sigma_0$
Lick, total field	5", 1/36	130 ^s	10.8 "	92	4	13.0	7.2	1/58.8	1/27.3	1.80 *	(3.4) †	0.74 $\times \sigma_0$
Lick, inner field	5", 1/36	130 ^s	10.6 "	43	4	7.5	4.8	1/26.1	1/15.6	1.11	(2.8) †	0.95 $\times \sigma_0$
Potsdam	8", 1/42	60 ^s	9.0 vis.	18	4	7.5	4.1	1/19.1	1/10.3		1.35	0.86 $\times \sigma_0$

In the case of the Lick field, which is by far the richest in stars, we have segregated an "inner field" of 43 stars extending to only 7.5 solar radii from the Sun. Despite the great superiority in the number of stars the relative weight of L for the whole Lick field of 92 stars is not substantially larger than the corresponding value for the Potsdam field of only 18 stars, namely 1.80 as against 1.35, and the standard deviation of the result is therefore only slightly smaller. This is chiefly due to the fact that the Lick Expedition employed the almost degenerate set of equations (7), whereas the Potsdam Expedition proceeded by the method described in section 2. Had the Potsdam field been symmetrical, it would have yielded a result whose weight would nearly have equalled that of the Lick field, namely 1.75 instead of 1.35. On the other hand, if the Lick field had been analysed with the aid of the non-degenerate equations (15), the relative weight of L would have been practically doubled; and this value would not have been greatly reduced, namely from 3.4 to 2.8, if the field had been limited to the inner region of 43 stars. The weight of L is therefore mainly due to the stars which are near the Sun. Hence all experiments must aim at observing as many stars as possible in the immediate vicinity of the Sun.

Incidentally, the Lick field is almost perfectly symmetrical, the reduction of w_L/w_0 owing to asymmetry amounting only to 3 per cent.

Let us now return to formula (23), namely

$$\delta L = -h \frac{1 - (\bar{x}\xi + \bar{y}\eta)}{1 - h(\xi^2 + \eta^2) - \frac{h(\bar{x}\eta - \bar{y}\xi)^2}{a - \bar{x}^2 - \bar{y}^2}} \delta S = -h' \delta S,$$

say, from which the following conclusions may be drawn: if 0".1 has been fixed as an upper bound for the standard deviation of L , then the standard deviation of S must not exceed

$$\frac{0''.1}{h'} 10^{-3},$$

if the solar radius, *i.e.* ($10^3''$), is taken as unit of length. If f denotes the focal length, this would mean that in the case of the three star fields in question (disregarding asymmetry)

$$\delta S = \delta f/f \leq 1.2 \times 10^{-5}; \quad 3.6 \times 10^{-6}; \quad 1 \times 10^{-5}$$

respectively; *i.e.* the focal lengths of the cameras would have to be known with an accuracy of

$$7.4 \times 10^{-2}; \quad 1.6 \times 10^{-2}; \quad 8.4 \times 10^{-2} \text{ mm.}$$

* The asymmetry terms have not been considered.

† The values in brackets were actually not obtained, because the scale correction S had not been separately determined.

respectively. It is obvious that such an accuracy cannot be guaranteed if, during the eclipse, the camera is swung round to a second star field from which the scale factor is to be derived. *The scale of each plate must be determined in exactly the same position in which the exposures on the eclipse are made.*

The necessity of this very exacting demand was actually experienced in the case of the plates taken in 1929 with the astrographic lens of aperture 8 inches and focal length 11.2 feet. During the eclipse the camera was turned towards the Hyades so that the scale factor might be ascertained. The eclipse field showed 30 stars at a mean distance of 6.7 solar radii. A scale factor $S^{(1)} = 0''.2259$ per solar radius was obtained in this way; on substituting in (15) it was found that the corresponding value of the light deflection was

$$L^{(1)} = 0''.77.$$

On the other hand, when the same observational data were treated with the scale factor

$$S^{(2)} = 0''.1875,$$

they yielded the value

$$L^{(2)} = 2''.45.$$

The difference between the two values of L is comparable with the value of L itself, and yet the difference between the values of S is only $0''.0384$, which is equivalent to a change in focal length of not more than $f/26000$ or 0.13 mm. Apparently the focal length of the camera changed by this small amount when the telescope was swung into the direction of the second star field.

In the Potsdam Expedition the scale correction was derived by exposing the plates to a fine grid without any change in the position of the camera. The determination was based on about 100 grid points and had an accuracy of $0''.0070$ per solar radius, which corresponds to 7×10^{-6} seconds of arc. This is even slightly less than the upper limit required, viz. 10^{-5} .

*The University,
St Andrews :
1943 November 22.*

THE SCHMIDT-CASSEGRAIN SYSTEMS AND THEIR APPLICATION TO ASTRONOMICAL PHOTOGRAPHY

E. H. Linfoot

(Received 1944 March 9)

Summary

The method of plate-diagram analysis, as developed and extended by Dr. C. R. Burch, is applied to a general survey of the Schmidt-Cassegrain systems. Explicit formulæ are obtained for the figuring strengths of the Baker astrographic cameras in terms of their two free parameters, for the monocentric Schmidt-Cassegrains and for the flat-fielded aplanats with both mirrors spherical; their distortion is also considered.

1. *Introduction.*—In the plate-diagram analysis * of an optical system S , we represent it as a sum of spherical surfaces (refracting or reflecting), each furnished with an anastigmatising plate, together with a “plate system”, which takes account (i) of the fact that the surfaces are not in general actually provided with such plates, (ii) of any asphericities which the surfaces may possess, (iii) of any aspherical plates which may be comprised in the system.

The physical difficulty that light could not, for example, reach the “missing” anastigmatising plate of a convex reflecting surface is removed when we Gauss-image such a plate through the part of the system which precedes it †; this procedure replaces



FIG. 1.

it by a plate, situated in “object space”, which has the same effect on the Seidel errors of the system S . Figurings on the surfaces themselves (if present) are imaged through in the same way into the object space. When the object is at infinity, which we suppose throughout to

be the case, the resulting set of plates, together with the entry pupil of the system, will reproduce the first three Seidel errors (spherical aberration, coma and astigmatism) of the actual system S . This set of plates, working in parallel light, we call the plate system belonging to S . If $\alpha_1, \alpha_2, \dots, \alpha_n$ are the strengths of the asphericities of the respective plates and $x_1, x_2, \dots, x_n$ their respective distances (positive against the light) from the entry pupil, then we can describe the plate system completely by means of the diagram of Fig. 1, which we call the plate-diagram of S . As unit of asphericity we choose that of the anastigmatising plate of a concave reflecting sphere of unit focal length.

The plate diagram yields simple expressions ‡ for the first three Seidel errors of S ; used in conjunction with Conrady's 3 : 1 rule and the Petzval equation, it therefore enables the field curvature to be simply calculated. To deal with distortion, however,

* C. R. Burch, *M.N.*, **102**, 159–165, 1942.

† The Gauss image of a plate is here defined as follows: its position and lateral size are found by Gaussian optics, while the retardation which it imparts to the wave-fronts passing through it at each point is to be the same after transformation as before.

‡ See *M.N.*, **103**, 216, 1943; §§ 9, 10 of this paper are vitiated by an algebraical error.

an extension of the analysis is required, because an anastigmatised sphere does not give distortionless imaging between object and image *planes*, but only between object and image *spheres* concentric with itself. Thus the sphere-plus-anastigmatising-plate elements into which we dissect *S* will make contributions to the distortion which must be added to those from the plate system to obtain the actual value of this error.

Dr. Burch has made this extension and has kindly allowed me to include the following account of his work in the present paper.

Fig. 2 shows one "element" of *S*, namely a spherical surface *PAA'* with anastigmatising plate. We shall call this the "element *A*" and *PAA'* the "surface *A*." The surface *A* refracts, or reflects, an arriving object-pencil from the (real or virtual) intermediate image *O* to form the intermediate image *I*; *N* and *N'* are the refractive indices

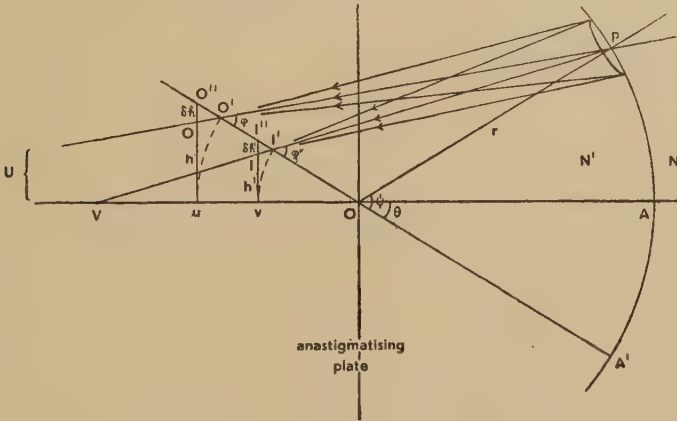


FIG. 2.

of the two spaces. The case of reflection* is covered by setting $N' = -N$. The plate can be located either in *N*-space or (with a suitable change of strength) in *N'*-space; one case can be obtained from the other by imaging the plate through the surface.

We define the "positions" of the intermediate images by means of the principal rays of the corresponding pencils. These are the rays which originally entered *S* through the centre of its entry pupil. *V* in Fig. 2 therefore agrees (to Gaussian accuracy) with the axial point of the Gauss image of the entry pupil in that part *S'* of *S* which forms the intermediate image *I*. Let *u*, *v* be the respective Gauss images of the axial point at infinity by the part *S''* of *S* which forms the intermediate object *O* and by *S'*. Their distances from *A* (positive in the figure) will be denoted by *u*, *v*. We call the planes, perpendicular to the optic axis, through *u*, *v* the "object and image planes for *A*"; the spheres, concentric with the surface *A*, through *u*, *v* we call the "object and image spheres for *A*". The respective intersections of the principal rays *PU*, *PV* with the object and image planes define the *image positions* *O*, *I* in these planes; their respective intersections with the object and image spheres define the *image positions* *O'*, *I'* in these spheres.

The line *O'C* we call the *auxiliary optic axis* of the surface *A* for the incident pencil under consideration; its angle θ with the main axis *CA* is a first order small quantity. Since a first order tilt of an anastigmatising plate about its centre leaves its Seidel properties unchanged, it follows that *I'* is on this auxiliary axis to third order accuracy. We can express this by saying that *the anastigmatic imaging effected by the element A is distortionless between object and image spheres*.

* This is the case which we shall require, but it is just as simple to discuss the more general case of arbitrary indices *N*, *N'*. When the surface *A* is a concave mirror, object and image are real in the figure; when *A* is a convex mirror both are virtual.

Since (to Gauss accuracy) the auxiliary axis CA' and the principal ray PIV meet on the image plane, they are transformed by the collineation (Gauss imaging back through S') into parallel lines through C_S and V_S ; V_S is the image of V in N_S -space, and so is the centre of the entry pupil, as remarked earlier, while the image of PIV is the entering principal ray. V_S is therefore the zero-point of the plate diagram of S . Expression (2) can now be written

$$-\frac{N'-N}{NN'}\omega^2(C_S V_S - A_S V_S) \frac{P_S A'_S}{P_S A_S} \cdot \frac{N_S}{2r} = -\frac{N'-N}{NN'} \cdot \frac{N_S}{r} \cdot C_S V_S \frac{\omega^2}{2}.$$

Here $-C_S V_S$ is the x -co-ordinate in the plate diagram of S of the image of the anastigmatising A-plate, while $\frac{N'-N}{NN'} \frac{N_S}{r}$ is the contribution from the A-surface to the Petzval

curvature of S ; we denote this by κ_A and obtain for the contribution of the A-element to the fractional distortion the expression

$$\frac{1}{2}\kappa_A x_A \omega^2.$$

Summing over all the surfaces of S , we have

$$\text{Fractional distortion of } S = \frac{1}{2}\omega^2 \sum_A \kappa_A x_A + \text{contribution from the plate system.}$$

The plate contribution to the distortion may be deduced from the fact that a plate of strength α at distance x from the entry pupil imposes an angular deviation $+\frac{1}{8}\omega^3 \alpha x^3$ on the principal ray.* This leads to a fractional distortion $-\frac{1}{8}\omega^2 \alpha x^3$. Hence finally,

$$\text{Fractional distortion of } S = -\frac{1}{8}\omega^2 \left(\sum_P \alpha x^3 - 4 \sum_A \kappa x \right), \quad (3)$$

where the first sum $\sum_P$ is over *all* the plates of the plate-system, the second $\sum_A$ over the imaged and sign-reversed Schmidt plates of the power-surfaces of the system. We write

$$\Delta = -\frac{1}{8} \left(\sum_P \alpha x^3 - 4 \sum_A \kappa x \right) \quad (4)$$

and call Δ the distortion-coefficient. A positive value of Δ indicates pincushion distortion.

Completed in this way, the plate-diagram method allows all the Seidel properties of an optical system to be expressed in a form which is particularly well adapted to the discussion of plate-mirror systems. In the present paper we use it to make a general survey of the Schmidt-Cassegrain systems (defined in section 2) and to obtain explicit formulæ for the figuring strengths of the general Baker cameras in terms of two parameters (substantially the component-spacings) which describe the general form of the system. Monocentric Schmidt-Cassegrains, Baker's preferred types A, B, C, D, and flat-fielded aplanats with both mirrors spherical are also discussed as special cases.

2. The Schmidt-Cassegrain Systems.—By a Schmidt-Cassegrain system we shall mean one consisting of a concave primary mirror M_1 , usually pierced, a convex secondary mirror M_2 , and an aspherical, nearly plane-parallel correcting plate P through which the light passes on its way to the primary. The subsequent path of the rays is similar to that in an ordinary Cassegrain telescope; that is to say, after reflection at the primary and secondary mirrors, they come to a focus near the pole A of the primary. See Fig. 4. One or both of the mirrors may be figured, but for ease of construction we try to keep the figurings on the mirrors as light as possible. In particular, systems in which one or both of the mirrors are spherical have a claim to preference on that account.

* See *M.N.*, 103, 216, 1943.

Suppose that f_1 is the paraxial focal length of the concave primary, f_2 that of the convex secondary, both being taken positively. Let $d > 0$ be the axial separation between the mirrors, x the distance of the plate in front of the primary. Let γ be the strength of the figuring on the plate, α that of the figuring on the primary, β that of the image in object space of the figuring on the secondary*, all in terms of Schmidt plate strength for a concave sphere of focal length 1 as unit. Then† the first three Seidel errors (spherical aberration, coma and astigmatism) of the system agree with those of the set

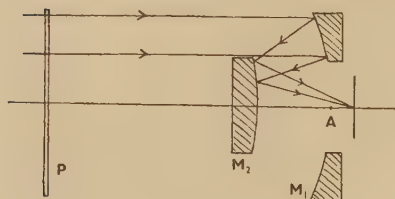


FIG. 4.

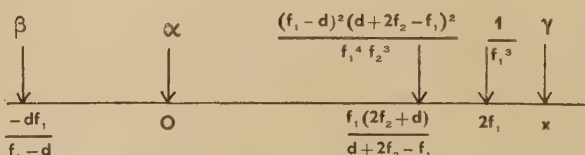


FIG. 5.

of aspheric plates, together with aperture stop, obtained by imaging through into object space the “missing” (*i.e.* sign-reversed) anastigmatising plates of the mirrors, the figurings on their surfaces, any actual aspheric surfaces which the system may contain, and the aperture stop of the system. The result for a Schmidt-Cassegrain system is shown in Fig. 5, in the case where the aperture stop is on the primary. It follows that the conditions for the vanishing of the first three Seidel errors of the system are respectively

$$\text{(Spherical aberration)} \quad \alpha + \beta + \gamma + \frac{(f_1 - d)^2(d + 2f_2 - f_1)^2}{f_1^4 f_2^3} - \frac{1}{f_1^3} = 0,$$

$$\text{(Coma)} \quad -\frac{df_1}{f_1 - d}\beta + x\gamma + \frac{f_1(2f_2 + d)}{d + 2f_2 - f_1} \frac{(f_1 - d)^2(d + 2f_2 - f_1)^2}{f_1^4 f_2^3} - \frac{2}{f_1^2} = 0,$$

$$\text{(Astigmatism)} \quad \left(\frac{df_1}{f_1 - d}\right)^2 \beta + x^2 \gamma + \frac{f_1^2(2f_2 + d)^2}{(d + 2f_2 - f_1)^2} \frac{(f_1 - d)^2(d + 2f_2 - f_1)^2}{f_1^4 f_2^3} - \frac{4}{f_1} = 0.$$

In systems where the spherical aberration and coma both vanish (and only such systems are of value as starting-points for the design of astrographic cameras), these conditions are independent of the position of the aperture stop.

Change of Notation.—We can simplify the notation by writing

$$A = \alpha f_1^3, \quad B = \beta f_1^3, \quad \Gamma = \gamma f_1^3; \quad \sigma = \frac{x}{f_1}, \quad q = 1 - \frac{d}{f_1}, \quad \xi = \frac{f_2}{f_1}.$$

Then the plate diagram becomes

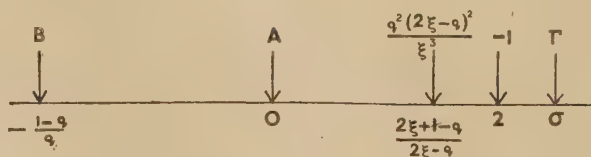


FIG. 6.

where now plate-strengths are measured in terms of the anastigmatising plate of M_1 as unit, and distances in terms of its focal length f_1 as unit. The conditions for the

* This notation has the advantage that α, β, γ are proportional to the actual number of fringes of asphericity at corresponding points on the three figured surfaces. The strength of the figuring on the secondary is $\beta f_1^4 / (f_1 - d)^4 = \beta / q^4$, where $q = 1 - d/f_1$.

† *M.N.*, 103, 213, 1943.

vanishing of the first three Seidel errors become

$$\text{(Spherical aberration)} \quad A + B + \Gamma + \frac{q^2(2\xi - q)^2}{\xi^2} = 1, \quad (5)$$

$$\text{(Coma)} \quad -\frac{1-q}{q}B + \sigma\Gamma + \frac{q^2(2\xi - q)(2\xi + 1 - q)}{\xi^3} = 2, \quad (6)$$

$$\text{(Astigmatism)} \quad \left(\frac{1-q}{q}\right)^2 B + \sigma^2\Gamma + \frac{q^2(2\xi + 1 - q)^2}{\xi^3} = 4. \quad (7)$$

These three equations specify all the Schmidt-Cassegrain anastigmats, while (5), (6) taken alone specify all the Schmidt-Cassegrain aplanats, in terms of six parameters each of which has a simple physical meaning. In fact,

A , B , Γ are the figuring *depths* on primary, secondary and plate, expressed in terms of parabolic correction of the primary as unit,

σ is the distance of the plate in front of the primary, expressed in terms of the paraxial focal length f_1 of the primary as unit,

ξ is the ratio f_2/f_1 of the paraxial focal length of the secondary to that of the primary,

q is the "minimum obstruction ratio" for the on-axis pencil*; alternatively, $1 - q$ is the separation between the mirrors, expressed in terms of f_1 as unit.

In these systems, the image is formed at a distance

$$\left(\frac{2\xi - q}{\xi - q}q - 1\right)f_1 \quad (8)$$

behind the pole of the primary, and the focal length

$$f = \frac{\xi}{\xi - q}f_1. \quad (9)$$

The Petzval curvature of the system

$$\frac{1}{\rho_P} = 2(D - C) = -\frac{1}{f_1} + \frac{1}{f_2} = \frac{1}{f_1} \frac{1 - \xi}{\xi} \quad (10)$$

and its astigmatism†

$$\frac{1}{\rho_T} - \frac{1}{\rho_S} = 4C = -\frac{1}{2}\Sigma\alpha x^2 = -\frac{1}{2f_1} \left[\left(\frac{1-q}{q}\right)^2 B + \sigma^2\Gamma + \frac{(2\xi + 1 - q)^2 q^2}{\xi^3} - 4 \right] \quad (11)$$

or, which is equivalent,

$$\frac{\text{Astigmatism}}{\text{Thin lens value}} = \frac{-\xi}{4(\xi - q)} \left[\left(\frac{1-q}{q}\right)^2 B + \sigma^2\Gamma + \frac{(2\xi + 1 - q)^2 q^2}{\xi^3} - 4 \right]. \quad (12)$$

Distortion in the Schmidt-Cassegrain Systems.—The contributions to the Petzval curvature from the two power-surfaces of a Schmidt-Cassegrain system are $\kappa_1 = -1/f_1$ from the primary mirror M_1 and $\kappa_2 = +1/f_2$ from the secondary M_2 . Applying equation (4) to the plate diagram of Fig. 6, we obtain for a Schmidt-Cassegrain with aperture

* That is, the obstruction ratio when the secondary is just big enough to receive the on-axis pencil. It will need to be slightly larger than this in order to receive the off-axis pencils also.

† *M.N.*, 103, 216-218, 1943.

stop on the primary

$$\begin{aligned}
 \Delta &= -\frac{1}{8} \sum_P \alpha x^3 + \frac{1}{2} \sum_A \kappa x \\
 &= -\frac{1}{8} \left\{ -B \left(\frac{1-q}{q} \right)^3 + \frac{(2\xi-q)^2 q^2}{\xi^3} \left(\frac{2\xi+1-q}{2\xi-q} \right)^3 - 8 + \Gamma \sigma^3 \right\} \\
 &\quad + \frac{1}{2} \left\{ (2f_1) \left(\frac{-1}{f_1} \right) + \frac{2\xi+1-q}{2\xi-q} f_1 \cdot \frac{1}{f_2} \right\} \\
 &= \frac{1}{8} B \left(\frac{1-q}{q} \right)^3 - \frac{(2\xi+1-q)^3 q^2}{8\xi^3(2\xi-q)} + 1 - \frac{1}{8} \Gamma \sigma^3 - 1 + \frac{2\xi+1-q}{2\xi(2\xi-q)} \\
 &= -\frac{1}{8} \left[\sigma^3 \Gamma - \left(\frac{1-q}{q} \right)^3 B \right] - \frac{2\xi+1-q}{2\xi(2\xi-q)} \left[\frac{(2\xi+1-q)^2 q^2}{4\xi^2} - 1 \right]. \quad (13)
 \end{aligned}$$

We next consider the effect on Δ of moving the aperture stop. Let ζf_1 be the change of position of the entry pupil caused by shift in stop-position; ζf_1 is measured positively against the light. Then the plate diagram is changed by having the origin shifted to the point formerly given by $x = \zeta f_1$. The new value of Δ is therefore

$$\begin{aligned}
 \Delta_\zeta &= -\frac{1}{8} \sum_P \alpha (x - \zeta f_1)^3 + \frac{1}{2} \sum_A \kappa (x - \zeta f_1) \\
 &= \Delta + \frac{3}{8} \zeta f_1 \sum_P \alpha x^2 - \frac{3}{8} \zeta^2 f_1^2 \sum_P \alpha x + \frac{1}{8} \zeta^3 f_1^3 \sum_P \alpha - \frac{1}{2} \zeta f_1 \sum_A \kappa \\
 &= \Delta + \zeta f_1 \left(\frac{3}{8} \sum_P \alpha x^2 - \frac{1}{2} \sum_A \kappa \right) \quad (14)
 \end{aligned}$$

when the spherical aberration $\sum_P \alpha$ and the coma $\sum_P \alpha x$ both vanish. If we are dealing with an anastigmat, $\sum_P \alpha x^2$ also vanishes, and if this anastigmat is flat-fielded $\sum_A \kappa = 0$. Equation (14) then reduces to

$$\Delta_\zeta = \Delta,$$

which shows that in a flat-fielded anastigmat, Δ is invariant under stop-shifting. In the more general case of a flat-fielded aplanat, $\sum_P \alpha x^2$ and $\sum_A \kappa$ will not usually vanish and (14) expresses the way in which distortion in the image plane varies under shifts of the aperture stop.

The Monocentric Schmidt-Cassegrains.—These are anastigmats in which the primary and secondary mirrors are concentric spheres and the plate passes through their common centre of curvature. The field is curved, being concentric with the mirrors, and this is the great drawback of the type. Against this must be set the fact that it yields better off-axis images with two spherical mirrors and one plate than do the flat-fielded, two-sphere aplanats (section 4) and is easier to make than the flat-fielded Baker anastigmats (section 3), in which at least two of the surfaces are aspherical.

The mirrors will be spherical if $A=0$, $B=0$ and concentric if $d=2f_1-2f_2$, i.e. $2\xi=1+q$. Inserting these values in (5), (6), (7) we obtain, as the equations satisfied by the monocentric Schmidt-Cassegrains,

$$\Gamma = 1 - \frac{8q^2}{(1+q)^3}, \quad \sigma\Gamma = 2 - \frac{16q^2}{(1+q)^3}, \quad \sigma^2\Gamma = 4 - \frac{32q^2}{(1+q)^3}.$$

The first two of these are equivalent to

$$\Gamma = 1 - \frac{8q^2}{(1+q)^3}, \quad \sigma = 2,$$

and the third is then redundant (corresponding to the fact that a monocentric aplanat is *eo ipso* an anastigmat). Thus the monocentric Schmidt-Cassegrains are included in our survey as the case

$$A=0, \quad B=0, \quad \Gamma=1-\frac{8q^2}{(1+q)^3}; \quad \sigma=2, \quad \xi=\frac{1+q}{2}. \quad (15)$$

The image-distance behind the primary is, by (8), $\frac{3q-1}{q-1}f_1$; the Petzval curvature is

$\frac{1}{f}=\frac{1-q}{1+q}\frac{1}{f_1}$. Since it is monocentric, the system gives distortionless imaging on its image sphere.

The Flat-fielded Anastigmats.—These are the systems first described by Baker.* The Seidel condition for flat-fieldedness in an anastigmat is that the Petzval curvature shall be zero. Thus the Baker cameras satisfy, in addition to (5), (6), (7), the condition $(1/f_2)-(1/f_1)=0$, i.e. $\xi=1$. Inserting this value, we obtain

$$A+B+\Gamma=1-q^2(2-q)^2, \quad (16)$$

$$-\frac{1-q}{q}B+\sigma\Gamma=2-q^2(2-q)(3-q), \quad (17)$$

$$\left(\frac{1-q}{q}\right)^2B+\sigma^2\Gamma=4-q^2(3-q)^2, \quad (18)$$

together with $\xi=1$, as the equations of the flat-fielded anastigmats. Their fractional distortion, which is independent of the stop position, is

$$\theta^2\Delta=\frac{1}{8}\theta^2\left[\left(\frac{1-q}{q}\right)^3B-\sigma^3\Gamma+\frac{3-q}{2-q}(4-q^2(3-q)^2)\right] \quad (19)$$

at angle θ off-axis.

Equations (16), (17), (18) enable us to express all the parameters of the system in terms of q and σ . First, elimination of Γ gives

$$\begin{aligned} \sigma A + \left(\sigma + \frac{1-q}{q}\right)B + q^2(2-q)[\sigma(2-q) - (3-q)] &= \sigma - 2 \\ -\frac{1-q}{q}\left(\sigma + \frac{1-q}{q}\right)B + q^2(3-q)[\sigma(2-q) - (3-q)] &= 2(\sigma - 2). \end{aligned}$$

Then, eliminating B , we obtain

$$\frac{1-q}{q}\sigma A + \left[\frac{1-q}{q}q^2(2-q) + q^2(3-q)\right][\sigma(2-q) - (3-q)] = (\sigma - 2)\frac{1+q}{q},$$

i.e.

$$(1-q)\sigma A + 2q^2[\sigma(2-q) - (3-q)] = (\sigma - 2)(1+q),$$

whence

$$A = \frac{(\sigma - 2)(1+q) - 2q^2[\sigma(2-q) - (3-q)]}{(1-q)\sigma}. \quad (20)$$

We also find

$$B = -\frac{q^2}{1-q} \frac{2(\sigma - 2) - q^2(3-q)[\sigma(2-q) - (3-q)]}{q\sigma + 1 - q}, \quad (21)$$

$$\Gamma = \frac{1}{\sigma} \left[2 - q^2(2-q)(3-q) - q \frac{2(\sigma - 2) - q^2(3-q)[\sigma(2-q) - (3-q)]}{q\sigma + 1 - q} \right]. \quad (22)$$

* *J. Amer. Phil. Soc.*, **82**, 339, 1940.

20), (21) and (22) express the figuring strengths in terms of the design-constants q (obstruction ratio) and σ (plate-primary separation). We choose these as the basic parameters of the system because they describe its general proportions and its vignetting properties, which demand first consideration in the practical design problem.

It is a practical convenience if the image lies behind the primary, or at any rate behind its front surface. By (8), the condition for this is

$$\frac{2-q}{1-q}q - 1 > 0 \quad \text{or} \quad q > 0.387.$$

Since, however, the strengths of the figurings on the mirrors increase with increasing q (see Tables I-IV below), it is sometimes better to choose q between 0.35 and 0.387. The image-surface is then still near enough to the primary to enable the photographic plates to be conveniently inserted and removed through a central hole in the primary.

3. *The Baker Camera Types A, B, C, D.*—The four one-parameter types singled out as specially valuable by Baker will next be derived as special cases of the general flat-fielded anastigmat.

Type A (short tube).

This is characterised by having the correcting plate at the focus of the primary, *i.e.* $\sigma = 1$. Then our equations become

$$\begin{aligned} A &= -1 - 2q, \\ B &= \frac{q^2}{1-q} [2 - q^2(3-q)], \\ \Gamma &= 2(1+q) - 2q^2(3-q). \end{aligned}$$

By equation (17), the distortion coefficient Δ satisfies the equation

$$\begin{aligned} 8\Delta &= \left(\frac{1-q}{q}\right)^3 B - \sigma^2 \Gamma + \frac{3-q}{2-q} [4 - q^2(3-q)^2] \\ &= \frac{2}{q} (1-q)(1-3q) + (1-q)^2(2-5q+q^2) + \frac{3-q}{2-q} [4 - q^2(3-q)^2]. \end{aligned}$$

In a system with $q = 0.3$ this gives $\Delta = -0.5694$, corresponding to a barrel distortion of 1/7 per cent. at 3° off-axis. Table I gives the numerical data of the system for selected values of q .*

TABLE I

q	A	B	Γ	σ	Δ	$f/f_1 = 1/(1-q)$
0.30	-1.6	0.2259	2.1140	1	-0.5694	1.4286
0.35	-1.7	0.3157	2.0508	1	-0.6339	1.5385
0.40	-1.8	0.4224	1.9680	1	-0.6606	1.6667
0.45	-1.9	0.5462	1.8673	1	-0.6605	1.8182
0.50	-2.0	0.6875	1.7500	1	-0.6406	2.0000
0.55	-2.1	0.8462	1.6178	1	-0.6056	2.2222
0.60	-2.2	1.0224	1.4720	1	-0.5589	2.5000
0.65	-2.3	1.2157	1.3143	1	-0.5032	2.8571
0.70	-2.4	1.4259	1.1460	1	-0.4406	3.3333

Type B (spherical secondary).

Setting $B = 0$ in the defining equations (16), (17), (18), we obtain as the equations of the general Baker B system

$$\begin{aligned} A + \Gamma &= 1 - q^2(2-q)^2, \\ \sigma \Gamma &= 2 - q^2(2-q)(3-q), \\ \sigma^2 \Gamma &= 4 - q^2(3-q)^2, \end{aligned}$$

* Tables I, II, III, IV are similar to tables given by Baker, *L.c.*, for the values $q = 0.1, 0.2, 0.3, 0.35, 0.4$. I have thought it worth while, however, to include also the values of the distortion coefficient Δ .

which we can write

$$A + \Gamma = P, \quad \sigma \Gamma = Q, \quad \sigma^2 \Gamma = R,$$

where *

$$P = 1 - q^2(2 - q)^2, \quad Q = 2 - q^2(2 - q)(3 - q), \quad R = 4 - q^2(3 - q)^2.$$

These have as solution

$$A = P - \frac{Q^2}{R} = -\frac{q^2(1 - q)^2}{R}, \quad B = 0, \quad \Gamma = \frac{Q^2}{R}; \quad \sigma = \frac{R}{Q}.$$

To find the distortion we set $B = 0$ in (19) and obtain

$$\Delta = \frac{1}{8} \left(-\sigma^3 \Gamma + \frac{3 - q}{2 - q} R \right) = \frac{1}{8} \left(-\frac{R^3}{Q^3} \frac{Q^2}{R} + \frac{3 - q}{2 - q} R \right) = -\frac{1}{4} \frac{R}{Q} \frac{1 - q}{2 - q}.$$

Thus all the Baker B systems suffer from barrel distortion. Its amount is very small; in the system with $q = 0.4$ (B5 in Baker's notation) $\Delta = -0.205$, corresponding to a distortion of 1/20 per cent. at 3° off-axis. Table II gives the numerical data of the system for selected values of q .

TABLE II

q	A	B	Γ	σ	Δ	$ff_1 = \frac{1}{1 - q}$
0.30	-0.0132	0	0.7531	2.1072	-0.2169	1.4286
.35	.0165	0	.6830	2.1441	.2112	1.5385
.40	.0197	0	.6101	2.1871	.2050	1.6667
.45	.0228	0	.5363	2.2368	.1984	1.8182
.50	.0256	0	.4631	2.2941	.1912	2.0000
.55	.0280	0	.3920	2.3604	.1831	2.2222
.60	.0299	0	.3243	2.4372	.1741	2.5000
.65	.0311	0	.2611	2.5268	.1638	2.8571
0.70	-0.0313	0	0.2032	2.6321	-0.1518	3.3333

Type C (spherical primary).

This is the case $A = 0$. Equations (16), (17), (18) now become

$$B + \Gamma = P, \quad -\left(\frac{1 - q}{q}\right)B + \sigma \Gamma = Q, \quad \left(\frac{1 - q}{q}\right)^2 B + \sigma^2 \Gamma = R,$$

which give

$$\left(\sigma + \frac{1 - q}{q}\right)B = \sigma P - Q, \quad -\frac{1 - q}{q} \left(\sigma + \frac{1 - q}{q}\right)B = \sigma Q - R,$$

* From their mode of definition, P , Q and R are seen to be the sign-reversed contributions of the two power-surfaces (excluding figurings) to the spherical aberration ΣA , the coma $\Sigma A\sigma$ and the astigmatism $\Sigma A\sigma^2$.

Table of Values of P , Q , R .

q	P	Q	R
0.30	0.73990	1.58690	3.34390
.35	.66649	1.46437	3.13974
.40	.59040	1.33440	2.91840
.45	.51349	1.19962	2.68324
.50	.43750	1.06250	2.43750
.55	.36399	0.92537	2.18424
.60	.29440	.79040	1.92640
.65	.22999	.65962	1.66674
0.70	0.17190	0.53490	1.40790

whence we have

$$-\frac{1-q}{q} = \frac{\sigma Q - R}{\sigma P - Q}, \quad \sigma = \frac{(1-q)Q + qR}{(1-q)P + qQ} = 1 + \frac{q(1-2q)}{(1-q)(1+2q-2q^2)},$$

$$B = \frac{\sigma P - Q}{\sigma + \frac{1-q}{q}} = -\frac{q^4(1-q)}{1+3q}, \quad \Gamma = P - B = P + \frac{q^4(1-q)}{1+3q}.$$

The general Baker C system is therefore given by

$$A=0, \quad B = -\frac{q^4(1-q)}{1+3q}, \quad \Gamma = 1 - q^2(2-q)^2 + \frac{q^4(1-q)}{1+3q}, \quad \sigma = 1 + \frac{q(1-2q)}{(1-q)(1+2q-2q^2)}.$$

Its distortion coefficient

$$\Delta = \frac{1}{8} \left[-(1-q)^4 \frac{q}{1+3q} - \sigma^3 \Gamma + \frac{3-q}{2-q} R \right].$$

Table III gives the numerical data of the system for selected values of q .

TABLE III

q	A	B	Γ	σ	Δ	$f/f_1 = \frac{1}{1-q}$
0.30	0	-0.0030	0.7429	2.1268	-0.2342	1.4286
.35	0	.0048	.6713	2.1684	.2290	1.5385
.40	0	.0070	.5974	2.2162	.2230	1.6667
.45	0	.0096	.5231	2.2709	.2161	1.8182
.50	0	.0125	.4500	2.3333	.2083	2.0000
.55	0	.0155	.3795	2.4047	.1994	2.2222
.60	0	.0185	.3129	2.4865	.1892	2.5000
.65	0	.0212	.2512	2.5808	.1774	2.8571
0.70	0	-0.0232	0.1951	2.6901	-0.1637	3.3333

Type D (distortion-free).

By (19), the condition for zero distortion is

$$\frac{1}{8} \left[\left(\frac{1-q}{q} \right)^3 B - \sigma^3 \Gamma + \frac{3-q}{2-q} R \right] = 0.$$

Hence the equations of the distortion-free Baker systems are

$$A + B + \Gamma = P, \quad -\frac{1-q}{q} B + \sigma \Gamma = Q,$$

$$\left(\frac{1-q}{q} \right)^2 B + \sigma^2 \Gamma = R, \quad -\left(\frac{1-q}{q} \right)^3 B + \sigma^3 \Gamma = \frac{3-q}{2-q} R.$$

These give

$$\sigma A + B \left(\sigma + \frac{1-q}{q} \right) = \sigma P - Q,$$

$$-\frac{1-q}{q} B \left(\sigma + \frac{1-q}{q} \right) = \sigma Q - R,$$

$$\left(\frac{1-q}{q} \right)^2 B \left(\sigma + \frac{1-q}{q} \right) = \sigma R - \frac{3-q}{2-q} R,$$

whence we have

$$-\frac{1-q}{q} = \frac{\sigma R - \frac{3-q}{2}R}{\sigma Q - R}; \quad (1-q)(\sigma Q - R) + q\left(\sigma R - R - \frac{R}{2-q}\right) = 0;$$
$$\sigma = \frac{2R}{(2-q)[(1-q)Q + R]}.$$

σ being determined, B, Γ and A are successively given by

$$B = \frac{q^2}{1-q} \frac{R - \sigma Q}{q\sigma + 1 - q}, \quad \Gamma = \frac{1}{\sigma} \left(Q + \frac{1-q}{q} B \right), \quad A = P - B - \Gamma.$$

Their explicit expressions in terms of q are too heavy to be of much value. Table IV gives the numerical data of the system for selected values of q .

TABLE IV

q	A	B	Γ	σ	Δ	$f/f_1 = \frac{1}{1-q}$
0.30	-0.2028	0.0399	0.9028	1.8609	0	1.4286
.35	.2451	.0612	.8504	1.8558	0	1.5385
.40	.2895	.0884	.7914	1.8537	0	1.6667
.45	.3360	.1221	.7274	1.8542	0	1.8182
.50	.3846	.1625	.6596	1.8571	0	2.0000
.55	.4354	.2102	.5892	1.8623	0	2.2222
.60	.4884	.2654	.5174	1.8696	0	2.5000
.65	.5436	.3284	.4452	1.8788	0	2.8571
0.70	-0.6011	0.3994	0.3736	1.8901	0	3.3333

4. *The Schmidt-Cassegrain Aplanats.**—These are high-definition flat-fielded systems, whose off-axis astigmatism is so small as to cause no observable effect on their image-size. This means, as was remarked by Slevogt, that they are for all practical purposes anastigmatic.

Setting $A=0$, $B=0$ in the general equations (5), (6) we obtain

$$\Gamma = 1 - \frac{q^2(2\xi - q)^2}{\xi^3}, \tag{19}$$

$$\sigma\Gamma = 2 - \frac{q^2(2\xi + 1 - q)(2\xi - q)}{\xi^3}. \tag{20}$$

By (10), (11)

$$\frac{1}{\rho_P} = -\frac{1}{f_1} + \frac{1}{f_2} = \frac{1}{f_1} \frac{1 - \xi}{\xi},$$

$$\begin{aligned} \frac{1}{\rho_T} - \frac{1}{\rho_S} &= 4C = \frac{1}{2f_1} \left[4 - \sigma^2\Gamma - \frac{(2\xi + 1 - q)^2 q^2}{\xi^3} \right] \\ &= \frac{1}{2f_1} \left[4 - \left(2 - \frac{q^2(2\xi + 1 - q)(2\xi - q)}{\xi^3} \right)^2 / \left(1 - \frac{q^2(2\xi - q)^2}{\xi^3} \right) - \frac{(2\xi + 1 - q)^2 q^2}{\xi^3} \right] \\ &= -\frac{1}{2\xi^3 f_1} \frac{1}{\xi^3 - q^2(2\xi - q)^2} \left([2\xi^3 - q^2(2\xi + 1 - q)(2\xi - q)]^2 \right. \\ &\quad \left. - [4\xi^3 - q^2(2\xi + 1 - q)^2][\xi^3 - q^2(2\xi - q)^2] \right) \\ &= -\frac{1}{2f_1} \frac{q^2(2\xi - q - 1)^2}{\xi^3 - q^2(2\xi - q)^2} \tag{21} \\ &= -\frac{q^2(2\xi - q - 1)^2}{2f_1 \xi^3 \Gamma}. \tag{22} \end{aligned}$$

* C. R. Burch, *M.N.*, 103, 159-165, 1943; E. H. Linfoot, *M.N.*, 103, 218-221, 1943; H. Slevogt, *Z. Instrk.*, Oct. 1942.

Here ρ_P , ρ_S , ρ_T are the radii of curvature of the Petzval surface, the sagittal focal surface, and the tangential focal surface of the system, taken as positive when the surfaces are concave towards the incident light.

(22) can be written in the equivalent form

$$\text{Astigmatism} = -\frac{\xi}{4(\xi - q)} \frac{q^2(2\xi - 1 - q)^2}{\xi^3 \Gamma} \times (\text{thin lens value}), \quad (23)$$

from which, remembering that ξ is nearly equal to 1 in all practical cases, we obtain the approximate formula

$$\text{Astigmatism} = -\frac{q^2(1 - q)}{4\Gamma} \times (\text{thin lens value}). \quad (23a)$$

The various types of field-flattening by controlled off-axis astigmatism are all covered by the equation

$$\frac{1}{\rho_T} - \frac{1}{\rho_S} = -\frac{\lambda}{\rho_P}, \quad (24)$$

where, in the Seidel approximation,

$$\lambda = 1 \text{ for a flat best field } \left(\frac{1}{\rho_T} + \frac{1}{\rho_S} = 0 \right),$$

$$\lambda = 2/3 \text{ for a flat tangential field } \left(\frac{1}{\rho_T} = 0 \right),$$

$\lambda = 2/5$ for the type of field-flattening, recommended by Conrady for high-definition systems, in which

$$\frac{1}{\rho_T} : \frac{1}{\rho_S} : \frac{1}{\rho_P} = 2 : 4 : 5.$$

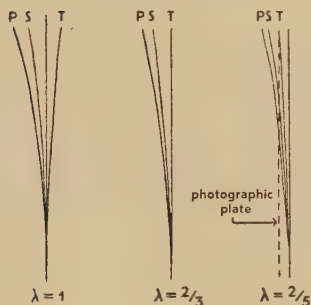


FIG. 7.

See Fig. 7. This "field-flattening equation" (24)

becomes, on substituting for $\frac{1}{\rho_P}$ and $\frac{1}{\rho_T} - \frac{1}{\rho_S}$,

$$-\frac{q^2(2\xi - q - 1)^2}{2f_1^3 \xi^3 \Gamma} = -\frac{\lambda(1 - \xi)}{f_1^3 \xi^3},$$

i.e.

$$\xi q^2(2\xi - q - 1)^2 = 2\xi^3 \Gamma \lambda(1 - \xi)$$

or

$$\xi q^2(2\xi - 1 - q)^2 = 2\lambda(1 - \xi)[\xi^3 - q^2(2\xi - q)^2]. \quad (25)$$

(25) gives the relation between ξ and q which will secure the desired type of field-flattening.* In practice, the value of ξ is generally between 0.95 and 1, and it is evident

* When $\lambda = 2/5$, this reduces to the corrected form of equation (17) of my previous paper, *M.N.*, 103, 219, 1943. In that paper equation (17), which purports to be the condition for a flat field, contains an algebraical error which vitiates section 9 from that point onward and results in the selection of an unsuitable (because not flat-fielded) system in section 9 and in the over-estimation of its error spreads in the best flat field by a factor 2 in section 10. This system is substantially a monocentric Schmidt-Cassegrain, and its off-axis errors other than primary field curvature are found to be exceedingly small even at a nominal aperture of $f/2$. But the field curvature causes error spreads, when the system is used with a flat photographic plate, amounting to 30" at the edge of a 4° diameter field, as compared with 40" for the Schwarzschild aplanat of the same effective aperture.

When (17) is corrected and the choice of system changed accordingly, a family of flat-fielded systems of performance much superior to that of the Schwarzschild aplanats is obtained, whose chief error is higher circular coma, traceable to the non-fulfilment of the exact sine condition.

that unless ξ is rather close to 1 the Petzval curvature $1/\rho_P$ will not be small and the astigmatism needed to flatten the field will spoil the definition.

In considering these systems of very small Petzval curvature, there is little to be obtained from *a priori* comparisons between the above-mentioned three types of astigmatic field-flattening. This is because the differences between them are comparable with the small empirical changes which are required in each design, after the pilot model has been constructed and its field explored with a microscope, in order to balance the image-errors over the required field. By these small empirical changes account is taken, without any *a priori* analysis, of:

(1) The contributions to the field curvature and to the astigmatism from the slight convexity which is superposed on the plate-figuring in order to minimise the colour error.*

(2) The higher aberrations. The primary astigmatism of the Schmidt-Cassegrain aplanats is very small, and in a system of effective aperture $F/2$ it is not detectable as such in the off-axis images, whose most important error is higher coma. In all the Schmidt-Cassegrain systems we can adjust the primary off-axis coma, without sensibly

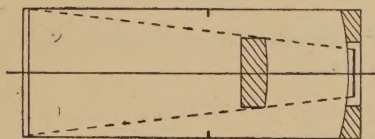


FIG. 8.

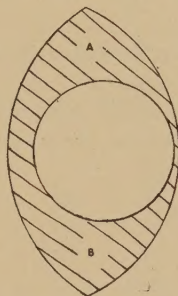


FIG. 9.

changing the other errors, by moving the plate along the axis. Thus our experimental modifications will include the adjusting of the plate-position along the axis so as to obtain the most satisfactory balance between higher coma and opposing primary coma over the desired field. At an aperture of $F/2$, the coma can be balanced over a field of 4° diameter sufficiently well to leave very little asymmetry in the images of an artificial star examined with a microscope.

(3) The effects of vignetting on the best-field curvature. When the field-diameter is fairly small, say 6° or less, we can avoid vignetting and secure a uniformly illuminated field, at the cost of some loss of aperture, by means of a stop placed approximately midway between the plate and the primary, as in Fig. 8. (We shall see later that this stop can be so placed as to annul the distortion of the system.) If, however, a larger field is desired, for example one of 10° diameter, this procedure becomes impracticable, partly because of the heavy sacrifice of aperture which it entails and still more because the areas sacrificed, which lie near the secondary plane of the pencil, are valuable regions of "small error slope", whose contributions help to build up a bright central nucleus in the images near the edge of the field. These regions are marked A, B in Fig. 9, which shows the appearance of the effective aperture. Now with an aperture of this form the best focus is determined mainly by the consideration that the image-patches contributed by the areas A and B must not disagree in position to any great extent,

* It is not difficult to take account of this plate-convexity in equations (19), (20) and (25), but in my experience nothing is gained by doing so; the theory loses its simplicity and we still need to determine the final design experimentally. It is usually advisable, however, to allow for it in calculating the final image-position for the pilot model. In Fig. 10, the image positions are shown without this allowance.

since such disagreement would produce a highly objectionable double-image effect. The position of best focus will therefore differ from that for a circular aperture, and the value of the effective field curvature will be changed accordingly.

Experience shows that, at apertures near $F/2$, a value of ξ very close to 1 and a little below it is needed to give a satisfactorily flat field. In designing the pilot model, therefore, we get closer to the final form by simply taking $\xi = 1$ than we should do by solving the three equations (19), (20), (25) for ξ , q and so deriving the explicit Seidel solution.

Setting $\xi = 1$ in (19), (20), we therefore obtain for the general two-sphere aplanats the practical design-equations

$$A=0, \quad B=0, \quad \Gamma = 1 - q^2(2-q)^2 = P;$$

$$\sigma = \frac{2 - q^2(2-q)(3-q)}{1 - q^2(2-q)^2} = \frac{Q}{P}, \quad \xi = 1.$$

When $0.4 \leq q \leq 0.75$ these give

$$\sigma = 2 + \frac{q^2(1-q)(2-q)}{\Gamma} \approx 2 + \frac{1}{6\Gamma},$$

an approximate formula connecting plate strength with plate-primary separation which is sometimes useful.

Table V gives the numerical values of the constants of the system, calculated from the above equations, for a selected set of values of q .

TABLE V

q	A	B	Γ	σ	$ff_1 = \frac{1}{1-q}$
0.30	0	0	0.7399	2.1447	1.4286
.35	0	0	.6665	2.1971	1.5385
.40	0	0	.5904	2.2602	1.6667
.45	0	0	.5135	2.3362	1.8182
.50	0	0	.4375	2.4286	2.0000
.55	0	0	.3640	2.5423	2.2222
.60	0	0	.2944	2.6848	2.5000
.65	0	0	.2300	2.8680	2.8571
0.70	0	0	0.1719	3.1117	3.3333

Distortion in the Two-Sphere Aplanats.—By (13), (14), the distortion in a two-sphere aplanat with stop on the primary is given by

$$\Delta = -\frac{1}{8}\sigma^3\Gamma - \frac{2\xi + 1 - q}{2\xi(2\xi - q)} \left[\frac{(2\xi + 1 - q)^2 q^2}{4\xi^2} - 1 \right], \quad (27)$$

while if the stop is at a distance ζf_1 in front of the primary the distortion-coefficient becomes

$$\Delta_\zeta = \Delta + \zeta f_1 \left(\frac{3}{8} \sum_P \alpha x^2 - \frac{1}{2} \sum_A \kappa \right). \quad (28)$$

(The image position is here defined by the intersection of the ray through the centre of the stop with the Gauss plane, that is to say by an ideal construction, since this ray is obstructed by the secondary in a Schmidt-Cassegrain system.) In most practical cases Δ is small; in the typical system *

$$\xi = 0.9205, \quad q = 0.3979, \quad \Gamma = 0.5773, \quad \sigma = 2.2249, \quad (29)$$

* This system is a solution of (19), (20) and (25) with $\lambda = 2/5$. An $F/2$ model which started from this solution finished with $\xi = 0.990$.

its value is -0.13 . This corresponds to a fractional distortion of 0.03 per cent. at 3° off-axis. Since, further, $f_1\left(\frac{3}{8}\sum_P ax^2 - \frac{1}{2}\sum_A \kappa\right)$ is small (its value is 0.12 in the above example) only a small change in Δ_ζ will be produced by any practicable shift of the

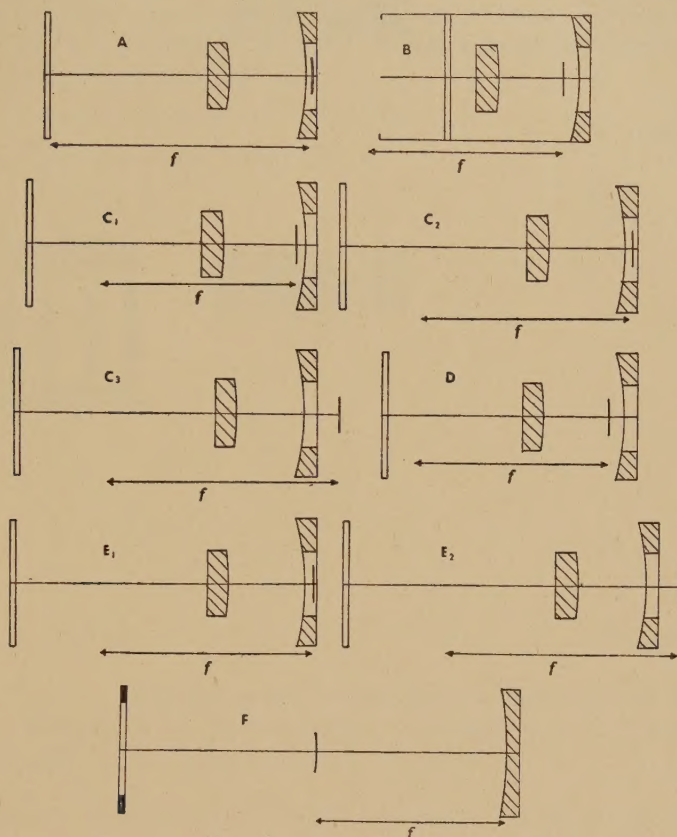


FIG. 10.

- (A) *Monocentric Schmidt-Cassegrain*, $q = 0.4$.
 (B) *Baker A system*, $q = 0.3$.
 (C_1), (C_2), (C_3) *Baker B systems*, $q = 0.35, 0.4, 0.45$. The Baker C systems are so close to these that the same diagrams can serve for both.
 (D) *Baker D system*, $q = 0.3$.
 (E_1), (E_2) *Two-sphere one-plate aplanats*, $q = 0.4, 0.45$.
 (F) *Ordinary Schmidt camera*, for comparison.

stop-position. The most satisfactory position of the stop is therefore governed mainly by considerations of vignetting. These more or less restrict the stop to lie between plate and primary, and point to a position approximately half-way between them as desirable in those cases where plate and primary are of equal diameter and a uniformly illuminated field is aimed at. This position corresponds to a value slightly larger than 1 for ζ . It is of some interest that in most cases the "distortion-free position of the stop", given by the equation

$$\zeta = \frac{\Delta}{\left(\frac{3}{8}\sum_P ax^2 - \frac{1}{2}\sum_A \kappa\right)f_1}, \quad (30)$$

is near this value. For example, in the system (29) the distortion-free position of the

stop is given by $\zeta = 0.9573$. However, even when the vignetting requirements are satisfactorily met by the value (30) of ζ , a deviation from this value will in general be found necessary to secure the least possible distortion in the physical images. This is because the intersection of the principal ray with the image plane does not exactly coincide with the mean centre of the physical off-axis image. Here, as elsewhere, an experimental investigation must supplement the Seidel theory if the best results are to be obtained.

Bristol University :

1944 March 9.